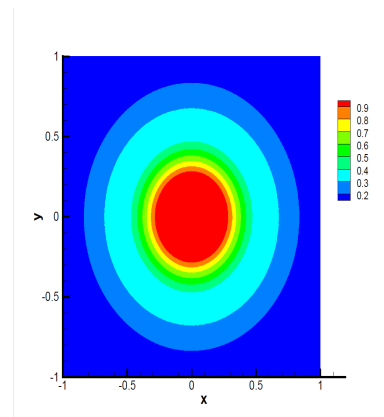
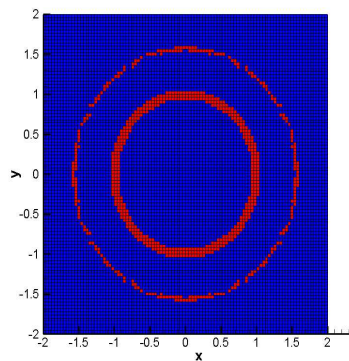
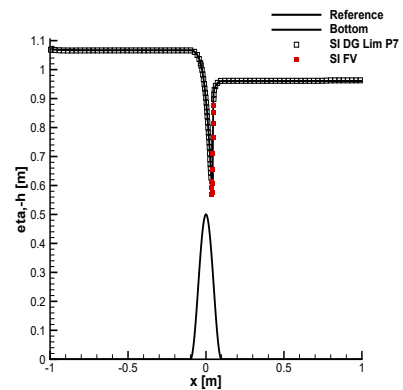
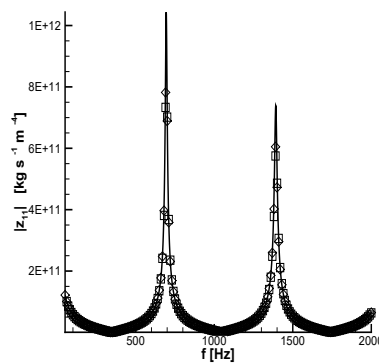


## Curriculum 1. Civil and Environmental Engineering

Matteo Ioriatti

# Semi-implicit schemes for compressible fluids in elastic pipes and a-posteriori sub-cell finite volume limiting techniques for semi-implicit Discontinuous Galerkin schemes for hyperbolic conservation laws on staggered meshes



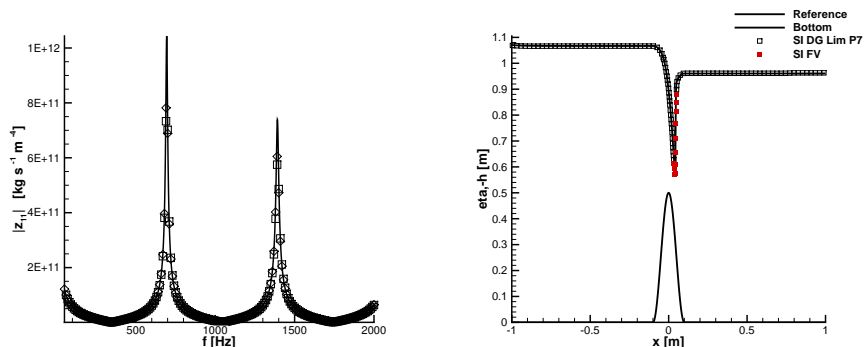
In this thesis partial differential equations have been solved using a new class of semi-implicit numerical methods on staggered meshes in order to analyse their properties and in order to show the potential applications of this family of scheme in the fields of computational fluid dynamics. In particular, two big topics are investigated. In the first part of the thesis, pressurized fluids in elastic pipes are investigated using staggered semi-implicit methods. In particular, the governing PDEs are first solved using finite volume schemes and successively they are integrated using arbitrary high order accurate Discontinuous Galerkin methods. Afterwards, in the second part we present a novel nonlinear a-posteriori stabilization technique for high order staggered semi-implicit Discontinuous Galerkin finite element schemes. This strategy is based on the combination of a-posteriori detection criteria with the use of staggered semi-implicit subcell finite volume schemes in those cells that have been detected as troubled.

The new limiter is applied to the shallow water equations and to the Euler equations of compressible gasdynamics. To the best knowledge of the author, this is the first time an a-posteriori subcell finite volume limiter has been proposed for high order discontinuous Galerkin schemes. In addition, this new tool is fundamental in order to carry out simulations for real world problems using high order Discontinuous Galerkin schemes, such as the treatment of wet and dry fronts in shallow water flows

**Matteo Ioriatti** was born in Trento on the 18th of July 1989. He has a master Degree in Environmental Engineering at the University of Trento with a dissertation in numerical analysis carried out under the supervision of Prof. Dr. Ing. Michael Dumbser. During his master thesis in 2014 he spent six months at the company Robert Bosch GmbH in the main department in Gerlingen-Schillerhöhe (Stuttgart). In particular, he investigated semi-implicit finite volume methods for compressible fluids in elastic tubes. During his PhD he continued the development of advanced semi-implicit schemes for computational fluid mechanics focusing the attention on pipe flow and on sub-cell finite volume limiters for semi-implicit high order Discontinuous Galerkin schemes on staggered cartesian meshes. He is a hard working person and in the free time he likes to swim and to ride the mountain bike

**Semi-implicit schemes for compressible fluids in elastic pipes and  
a-posteriori sub-cell finite volume limiting techniques for  
semi-implicit Discontinuous Galerkin schemes for hyperbolic  
conservation laws on staggered meshes**

Matteo Ioriatti



UNIVERSITÀ DEGLI STUDI DI TRENTO

Dipartimento di Ingegneria Civile, Ambientale e Meccanica

2018

Credits of the cover image: Matteo Ioriatti



Except where otherwise noted, contents on this book are licensed under a Creative  
Common Attribution - Non Commercial - No Derivatives  
4.0 International License

University of Trento  
Doctoral School in Civil, Environmental and Mechanical Engineering  
<http://web.unitn.it/en/dricam>  
Via Mesiano 77, I-38123 Trento  
Tel. +39 0461 282670 / 2611 - [dicamphd@unitn.it](mailto:dicamphd@unitn.it)



Doctoral thesis in **Civil, Environmental and Mechanical Engineering, XXXI cycle.**

Department of Civil, Environmental and Mechanical Engineering,  
**University of Trento.**

Academic year **2018/2019.**

Internal supervisor                      **Prof. Dr.-Ing. Michael Dumbser**

External supervisor                      **Dr. Raphaël Loubère**  
*University of Toulouse, France*

Examination commission              **Prof. Dr. Claus-Dieter Munz**  
*University of Stuttgart, Germany*  
**Prof. Stéphane Louis Clain**  
*University of Minho, Portugal*  
**Prof. Vincenzo Casulli**  
*University of Trento, Italy*

University of Trento  
Trento, Italy  
2018

Alla mia famiglia e ai miei amici

*“per alta virtus it”*

*Seneca*

## Preface

The material collected in this thesis has been developed during the three-years Doctoral program in Environmental Engineering at the Department of Civil, Environmental and Mechanical (DICAM) Engineering of the University of Trento.

The doctoral study grant was funded by DICAM and by the German company *Robert Bosch GmbH*. Most of the research has been carried out in Trento supervised by Prof. Dr.-Ing. Michael Dumbser. During 2017 I spent three months in Toulouse at the University Paul Sabatier under the supervision of Dr. Raphaël Loubère. In addition, in September 2018 I spent four weeks in Stuttgart at the *Institut für Aerodynamik und Gasdynamic*, where I participated at the *DROPIT* summer school.

In this thesis partial differential equations have been solved using a new class of semi-implicit numerical methods on staggered meshes in order to analyse their properties and in order to show the potential applications of this family of scheme in the fields of computational fluid dynamics. In particular, two big topics are investigated. In the first part of the thesis, pressurized fluids in elastic pipes are investigated using staggered semi-implicit methods. In particular, the governing PDEs are first solved using finite volume schemes and successively they are integrated using arbitrary high order accurate Discontinuous Galerkin methods. Afterwards, in the second part we present a novel nonlinear *a posteriori* stabilization technique for high order staggered semi-implicit Discontinuous Galerkin finite element schemes. This strategy is based on the combination of a *a posteriori* detection criteria with the use of staggered semi-implicit subcell finite volume schemes in those cells that have been detected as troubled. The new limiter is applied to the shallow water equations and to the Euler equations of compressible gasdynamics. To the best knowledge of the author, this is the first time an *a posteriori* subcell finite volume limiter has been proposed for high order discontinuous Galerkin schemes. In addition, this new tool is fundamental in order to carry out simulations for real world problems using high order Discontinuous Galerkin schemes, such as the treatment of wet and dry fronts in shallow water flows.

Trento, September 2018

Matteo Ioriatti

## Acknowledgments

I want to thank my supervisor Michael Dumbser for the time that he dedicated to my work during the doctoral program. Moreover, I wish to thank him for the possibility that he gave me to work in his group of research and for all the things that I learned. In addition, I am grateful to him for the experiences that I have made in the places that I visited during my PhD studies.

I wish also to thank the colleagues of the Laboratory of Applied Mathematics for their friendship and for the help given over these years. In particular special thanks are for Maurizio Tavelli and Francesco Fambri because we shared the office space and they gave me a lot of helpful advice from the first day of work.

I thank the professors and the secretaries of the department respectively for all the things that I learned and for the support for the administrative procedures. Then, I thank the company *Robert Bosch GmbH* because they co-funded my grant. I am also grateful to Raphaël Loubère for the collaboration of three months done with him during 2017 in Toulouse. In addition, I thank Prof. Dr. Claus-Dieter Munz and all the members of his research team for the time that I spent in Stuttgart at the *Institut für Aerodynamik und Gasdynamik* during the *DROPIT* summer school in September 2018.

Finally, I thank my family for all the sacrifices that they have done in order to make me study engineering. Last but not least, I thank my friends and the people that I met during my studies for all the good moments spent together.



# Contents

<b>Abbreviations</b>	<b>xi</b>
<b>Abstract</b>	<b>xii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Explicit schemes . . . . .	1
1.1.1 Explicit finite volume schemes . . . . .	2
1.1.2 Explicit Discontinuous Galerkin schemes . . . . .	5
1.2 Implicit and semi-implicit schemes . . . . .	7
1.3 Targets of the thesis . . . . .	9
1.4 Structure of the thesis . . . . .	9
<b>2 Explicit and semi-implicit FV schemes for compressible flows in elastic pipes</b>	<b>11</b>
2.1 Mathematical models and numerical methods for pipe flow . . . . .	12
2.1.1 Explicit path-conservative finite volume schemes for non-conservative hyperbolic systems . . . . .	12
2.1.2 Semi-implicit finite volume schemes . . . . .	15
2.1.3 Unsteady friction models . . . . .	21
2.1.4 A new efficient approximation of the Zielke convolution integral . . . . .	22
2.2 Numerical results . . . . .	23
2.2.1 Womersley test . . . . .	24
2.2.2 Impedance matrix for weakly compressible flow in the frequency domain . . . . .	24
2.2.3 Water hammer . . . . .	39
2.2.4 Efficiency analysis and computational times . . . . .	39
2.3 Conclusions about explicit vs. semi-implicit schemes for pipe flow . . . . .	55
<b>3 Staggered SI DG for axially symmetric viscous compressible flows in elastic tubes</b>	<b>59</b>
3.1 Staggered semi-implicit DG scheme applied to the 1D model . . . . .	60
3.1.1 Governing equations of the 1D model . . . . .	60

3.1.2	Staggered grids . . . . .	61
3.1.3	Basis functions . . . . .	61
3.1.4	DG scheme for the 1D model . . . . .	63
3.1.5	Non-linear convective terms for the 1D model . . . . .	67
3.2	Staggered semi-implicit DG scheme applied to the $2D_{xr}$ model . . . .	68
3.2.1	Governing equations of the $2D_{xr}$ model . . . . .	68
3.2.2	Computational grids and basis functions . . . . .	68
3.2.3	DG scheme for the $2D_{xr}$ model - Integral continuity equation .	69
3.2.4	DG scheme for the $2D_{xr}$ model - Momentum equation - Part A	70
3.2.5	DG scheme for the $2D_{xr}$ model - Momentum equation - Part B	74
3.3	Numerical tests . . . . .	76
3.3.1	Steady flow in an elastic pipe and numerical convergence study	76
3.3.2	Womersley profile . . . . .	78
3.3.3	Impedance matrix . . . . .	80
3.3.4	Blasius boundary layer . . . . .	82
3.4	Conclusions on staggered semi-implicit DG schemes for pipe flow . . .	85
<b>4</b>	<b>A posteriori sub-cell FV limiting of staggered SIDG for SWE</b>	<b>87</b>
4.1	Sub-cell finite volume limiting of staggered semi-implicit DG schemes in 1D . . . . .	88
4.1.1	Governing equations of the 1D model . . . . .	88
4.1.2	Unlimited staggered semi-implicit DG scheme for the 1D shal- low water equations . . . . .	88
4.1.3	A sub-cell formulation for the semi-implicit finite volume method for the 1D shallow water equations . . . . .	92
4.1.4	MOOD algorithm and detection criteria - 1D case . . . . .	95
4.1.5	Sub-cell limiting of the semi-implicit DG scheme for the 1D shallow water equations . . . . .	98
4.1.6	A high-resolution wetting and drying algorithm for high order staggered semi-implicit DG in 1D . . . . .	105
4.2	Sub-cell finite volume limiting of staggered semi-implicit DG schemes in 2D . . . . .	106
4.2.1	Governing equations of the 2D model . . . . .	106
4.2.2	Unlimited staggered semi-implicit DG method for the 2D shal- low water equations . . . . .	107
4.2.3	Sub-cell formulation for the finite volume method for the 2D shallow water equations . . . . .	111
4.2.4	MOOD algorithm and detection criteria - 2D case . . . . .	113

4.2.5	Sub-cell limiting of the semi-implicit DG scheme for the 2D shallow water equations . . . . .	115
4.3	Numerical tests for the one-dimensional model . . . . .	123
4.3.1	Well-balanced property for the 1D case . . . . .	123
4.3.2	One-dimensional Riemann problems . . . . .	126
4.3.3	Trans-critical flow over a bump . . . . .	132
4.3.4	Oscillating lake in a parabolic bowl . . . . .	132
4.4	Numerical tests for the two-dimensional model . . . . .	133
4.4.1	Well-balanced property for the 2D case . . . . .	133
4.4.2	Circular dambreak . . . . .	136
4.5	Conclusions about SIDG with subcell FV limiter for SWE . . . . .	140
<b>5</b>	<b>A posteriori sub-cell FV limiting of staggered SIDG for the Euler equations</b>	<b>141</b>
5.1	Staggered SIDG with limiter for the 1D Euler equation . . . . .	142
5.1.1	Governing equations of the 1D model . . . . .	142
5.1.2	Unlimited staggered semi-implicit DG scheme for the 1D Euler equations . . . . .	143
5.1.3	A sub-cell formulation for the semi-implicit finite volume method for the 1D Euler equations . . . . .	149
5.1.4	MOOD algorithm and detection criteria - 1D case . . . . .	152
5.1.5	Lmiting of the SI DG scheme for the 1D Euler equations . . . . .	154
5.2	Numerical tests for the one-dimensional model . . . . .	163
5.2.1	Advection of a smooth density wave in 1D . . . . .	163
5.2.2	One-dimensional Riemann problems . . . . .	163
5.3	Numerical tests for the two-dimensional model . . . . .	173
5.3.1	Isentropic vortex . . . . .	173
5.3.2	Advection of a smooth density bell in 2D . . . . .	175
5.3.3	Smooth two-dimensional acoustic wave propagation . . . . .	176
5.3.4	Circular explosion . . . . .	178
5.3.5	Two dimensional Riemann problems . . . . .	178
5.4	Conclusions about SIDG with subcell FV limiter for the Euler equations	180
<b>6</b>	<b>Conclusions and Outlook</b>	<b>187</b>
<b>A</b>	<b>Appendices</b>	<b>189</b>
A.1	Tensors for the staggered SI schemes for the 1D SWE . . . . .	189
A.1.1	Matrices and tensors for the 1D semi-implicit DG method . . . . .	189



A.1.2	Matrices and tensors for the 1D semi-implicit sub-cell FV method for P=2 . . . . .	190
A.2	Tensors for the staggered SI schemes for the 2D SWE . . . . .	190
A.2.1	Matrices and tensors for the 2D semi-implicit DG method . . .	190
A.2.2	Matrices and tensors for the 2D semi-implicit sub-cell FV method for P=1 . . . . .	192
A.3	Matrices and tensors for the 1D semi-implicit schemes for Euler equation	194
A.3.1	Matrices and tensors for the 1D semi-implicit DG method . . .	194
A.3.2	Matrices and tensors for the 1D semi-implicit sub-cell FV method for P=2 for Euler equation . . . . .	194
A.3.3	Tensors for the limited 1D semi-implicit DG method for Euler equation . . . . .	195
<b>Bibliography</b>		<b>197</b>
<b>List of Tables</b>		<b>210</b>
<b>List of Figures</b>		<b>212</b>

## Symbols

## Abbreviations

1D	One-dimensional
2D	Two-dimensional
ADER	Arbitrary high order scheme using derivatives
CG	Conjugate Gradient method
CI	Convolution integral friction model
CFL	Courant-Friedrichs-Levy number
DG	Discontinuous Galerkin method
ENO	Essentially non-oscillatory
FV	Finite Volume method
GMRES	GMRES: A Generalized Minimal Residual Algorithm for Solving Nonsymmetric linear systems
IA	Instantaneous acceleration friction model
IC	Initial condition
MOOD	Multi-dimensional optimal order detection
ODE	Ordinary differential equation
PDE	Partial differential equation
RK	Runge-Kutta method
RKDG	Runge-Kutta Discontinuous Galerkin method
RP	Riemann Problem
RS	Riemann Solver
SI	Semi-implicit method
SIDG	Semi-implicit Discontinuous Galerkin method
SWE	Shallow water equations
TVD	Total variation diminishing scheme

## Abstract

In the present work we solve systems of partial differential equations (PDE) for hyperbolic conservation laws using semi-implicit numerical methods on *staggered* meshes applied both to the class of finite volume (FV) and both to the family of high order Discontinuous Galerkin (DG) finite elements schemes. In particular, we want to show that these new semi-implicit schemes can be applied in several fields of applied sciences, such as in geophysical flows and compressible fluids in compliant tubes.

Inside this thesis we distinguish two big parts. First, we consider staggered semi-implicit schemes for compressible viscous fluids flowing in elastic pipes. This topic is very important in several practical applications of civil, environmental, industrial and biomedical engineering. Here, we analyse the accuracy and the computational efficiency of fully explicit and semi-implicit 1D and 2D finite volume schemes for the simulation of highly unsteady viscous compressible flows in laminar regime in axially symmetric rigid and elastic pipes. We consider two families of differential models that can be used to predict the pressure and velocity distribution along the tube. One is the so called  $2D_{xr}$  PDE model which is derived from the full compressible Navier-Stokes equations under the assumptions of a hydrostatic pressure and an axially symmetric geometry. The second family is a simple 1D non-conservative PDE system based on the cross-sectionally averaged version of the Navier-Stokes equations in cylindrical coordinates. In this last case, the use of a simple steady friction model is not enough to simulate the wall friction phenomena in highly transient regime. As a consequence the wall friction model has to be *frequency dependent* and, following previous studies present in the literature, we consider the classes of convolution integral (CI) models and instantaneous acceleration (IA) models.

We carry out a rather complete analysis of the previously-mentioned methods for the simulation of flows characterized by fast transient regime in rigid and compliant tubes. The numerical results show that the convolution integral models are clearly better than instantaneous acceleration models concerning accuracy. Moreover, for CI models, instead of computing the convolution integrals, which is very time- and memory-consuming, we express these methods via a set of additional ODEs for appropriate auxiliary variables. This trick improves the computational efficiency of these methods substantially, since it avoids the direct computation of the convolution integral. In addition, semi-implicit finite volume methods are significantly superior to classical explicit finite volume schemes in terms of computational efficiency, however, providing the same level of accuracy.

We then proceed by extending the finite volume discretization of the 1D and  $2D_{xr}$  PDE models to arbitrary high-order of accuracy in space introducing a new SIDG scheme

on staggered meshes. Both models include the effects of the viscosity and of the wall motion. The nonlinear convective terms are discretized explicitly by using a classical RKDG scheme of arbitrary high-order of accuracy in space and third order of accuracy in time. The continuity equation is integrated over the elements that belong to the main grid, while the momentum equation is integrated over the control volumes of the edge-based staggered dual grid. Inserting the discrete momentum equation into the discrete continuity leads to a mildly nonlinear algebraic system for the degrees of freedom of the pressure, which is solved by using the (nested) Newton method of Brugnano, Casulli and Zanolli. We use the  $\theta$ -method in order to get second order of accuracy in time for the implicit part of the scheme. In addition, the schemes have to obey only a mild CFL condition based on the fluid velocity and not based on the sound speed; consequently these schemes work also in the low Mach number regime and even in the incompressible limit of the Navier-Stokes equations. This is a very important property, which is the so-called asymptotic preserving (AP) property of the scheme. We carry out several numerical tests in order to validate this novel family of numerical methods against available exact solutions and experimental data. We also report numerical convergence tables in order to show that the new schemes indeed achieve high order of accuracy in space.

In the second part of the thesis, we present a new class of *a posteriori* sub-cell finite volume limiters for spatially high order accurate semi-implicit discontinuous Galerkin schemes on staggered Cartesian grids for the solution of the 1D and 2D shallow water equations (SWE) and of the Euler equations both expressed in conservative form. Here, the starting point is the unlimited arbitrary high order accurate staggered SIDG scheme proposed by Dumbser and Casulli (2013). For this method, the mass conservation equation and the momentum equations are integrated using a discontinuous finite element strategy on staggered control volumes, where the discrete free surface elevation is defined on the main grid and the discrete momentum is defined on edge-based staggered dual control volumes. According to the semi-implicit approach, pressure terms are discretized implicitly, while the nonlinear convective terms are discretized explicitly. Inserting the momentum equations into the discrete continuity equation leads to a well conditioned block diagonal linear system for the free surface elevation which can be efficiently solved with modern iterative methods.

Furthermore, the staggered SIDG is also extended to the Euler equations of compressible gasdynamics. Here, the governing PDE are rewritten using a flux vector splitting technique. The convective terms are updated using an explicit Runge-Kutta DG integrator. Then, the discrete momentum equation, which is integrated again on the dual grid, is coupled with the discrete energy equation that is discretized on the control volumes of the main grid. The pressure is efficiently obtained solving a linear system combined with an iterative Picard iteration procedure.

However, according to Godunov's theorem, any unlimited high order scheme inevitably produces spurious oscillations in the vicinity of discontinuities and strong gradients.

Therefore, in this PhD thesis the family of a posteriori sub-cell finite volume limiters recently introduced by Dumbser et al. (2014) for explicit DG schemes is also extended *for the first time ever* to semi-implicit time discretizations.

At time  $t^n$  the unlimited DG scheme is run in order to produce a so-called candidate solution for time  $t^{n+1}$ . Later on, we find the cells characterized by a non-admissible candidate solution using physical and numerical detection criteria based on the positivity of the solution, the absence of floating point errors and the use of a relaxed discrete maximum principle (DMP) according to the MOOD strategy of Clain, Loubère and Diot (2013).

In all the cells marked as troubled control volumes a lower order but more robust semi-implicit finite volume (FV) method is then applied on a sub-grid composed of  $2P + 1$  cells, where  $P$  denotes the polynomial degree used for approximating the discrete solution within the DG scheme. Then, after having identified the troubled cells, the linear system for the new pressure terms is assembled and solved again, where unlimited cells use the high order semi-implicit DG scheme and limited cells are evolved via the more robust finite volume method on the subgrid. Finally, from the subcell finite volume averages a higher order DG polynomial is reconstructed and the scheme proceeds with the next time step.

We apply the new semi-implicit staggered DG method with a posteriori subcell FV limiter to classical benchmarks such as Riemann problems in 1D and circular explosion problems in 2D with shock waves, showing that the new subcell limiter resolve shocks successfully and accurately without producing spurious oscillations. Moreover, in the benchmarks characterized by smooth solutions the detector does not find any troubled cells. As consequence, the limiter is not activated and the method corresponds to the unlimited staggered semi-implicit DG scheme. For SWE we carry out additional numerical tests in order to show the well-balanced property of the method and that it is also able to deal with wet and dry fronts properly.



# 1 Introduction

Many problems and phenomena investigated by science and engineering can be expressed by the principle of conservation of certain physical properties, such as mass, momentum and total energy. Consequently, the behaviour of a physical system is governed by conservation laws, which are mathematically classified as partial differential equations (PDEs), since the physical quantities in general depend on space and time.

In particular, the phenomena of the fundamental branches of continuum mechanics are described by PDEs, but for this family of equations, exact solutions are available only in very simple situations. Over the last fifty years, great advances have been made in the field of numerical analysis and the approximation of PDEs has a fundamental importance in engineering and applied sciences.

In the following, without claiming to be exhaustive, we briefly recall the most widespread strategies adopted in order to approximate PDEs numerically. For every class, we state the advantages and the disadvantages, in particular, focusing the attention on semi-implicit schemes, since this is the family of methods that has been mainly investigated in this thesis. For the sake of clarity and in order to ease notation, in this introduction we consider only rather simple examples of PDEs in one space dimension, in to highlight the benefits and drawbacks of the different methods.

## 1.1 Explicit schemes

First, we consider explicit methods, which are one of the most widespread strategies for solving non-linear hyperbolic time-dependent systems of conservation laws. A system of  $m$  conservation laws is given as follows

$$\frac{\partial \mathbf{Q}}{\partial t} + \frac{\partial}{\partial x} \mathbf{F}(\mathbf{Q}) = 0, \quad \mathbf{Q} \in \Omega_Q \subset \mathbb{R}^m, \mathbf{F}(\mathbf{Q}) \in \mathbb{R}^m, \quad (1.1)$$

where  $\mathbf{Q}$  is the vector of the conservative variables and  $\mathbf{F}(\mathbf{Q})$  is the flux vector. The system above is rewritten in the quasi-linear form

$$\frac{\partial \mathbf{Q}}{\partial t} + \mathbf{A}(\mathbf{Q}) \frac{\partial \mathbf{Q}}{\partial x} = 0, \quad \text{with} \quad \mathbf{A}(\mathbf{Q}) = \frac{\partial \mathbf{F}}{\partial \mathbf{Q}}, \quad (1.2)$$

and where  $\mathbf{A}(\mathbf{Q})$  is the so-called Jacobian matrix of the flux  $\mathbf{F}(\mathbf{Q})$ ; the system is called *hyperbolic* if the matrix  $\mathbf{A}$  has  $m$  real eigenvalues and  $m$  linearly independent eigenvectors. We distinguish three categories of explicit methods. The first is the family of explicit finite difference schemes (FD), where data are represented by point values. The second one are the finite volume (FV) methods, where data are represented by cell averages inside each cell. The third type of explicit schemes considered here is the so-called Discontinuous Galerkin (DG) finite element method (FEM), characterized by high order polynomial data inside each control volume, with a discontinuity (jump) at the cell interfaces.

### 1.1.1 Explicit finite volume schemes

Here, we explain shortly the main features and the numerical background of explicit finite volume schemes; for an exhaustive treatment of these methods see [148]. We denote the computational domain in space by  $\Omega = [x_L, x_R]$ , discretized by  $N_x$  equidistant cells, indicated as  $T_i = [x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}]$ , with constant length  $\Delta x = \frac{x_R - x_L}{N_x}$ . We introduce also a temporal control volume  $T_n = [t^n, t^{n+1}]$ . Hence the elementary space-time control volume is denoted as  $T_i \times T_n$ . We now integrate eq. (1.1) over  $T_i \times T_n$  and obtain

$$\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \mathbf{Q}(x, t^{n+1}) dx = \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \mathbf{Q}(x, t^n) dx + \int_{t^n}^{t^{n+1}} \mathbf{F}(\mathbf{Q}(x_{i-\frac{1}{2}}, t)) dt - \int_{t^n}^{t^{n+1}} \mathbf{F}(\mathbf{Q}(x_{i+\frac{1}{2}}, t)) dt. \quad (1.3)$$

Defining the cell averages at the time levels  $t^n$  by

$$\mathbf{Q}_i^n = \frac{1}{\Delta x} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \mathbf{Q}(x, t^n) dx \quad (1.4)$$

and, successively, the time averaged fluxes at the interfaces  $x_{i\pm\frac{1}{2}}$  as

$$\mathbf{F}_{i\pm\frac{1}{2}}^n = \frac{1}{\Delta t} \int_{t^n}^{t^{n+1}} \mathbf{F}(\mathbf{Q}(x_{i\pm\frac{1}{2}}, t)) dt, \quad (1.5)$$

we end up with the following reformulation

$$\mathbf{Q}_i^{n+1} = \mathbf{Q}_i^n - \frac{\Delta t}{\Delta x} (\mathbf{F}_{i+\frac{1}{2}}^n - \mathbf{F}_{i-\frac{1}{2}}^n), \quad (1.6)$$



which is the base in order to build an approximate solution of the system (1.1). In the FV framework, the key factor is the choice of the numerical flux at the cell interfaces. The most appropriate option is the Godunov flux, which is based on the exact solution of the Riemann problem at the cell interface [45, 78, 148]. Since this procedure is also the most difficult and time-consuming one, several approximate Riemann solvers (RS) have been proposed in literature. A first example is the Riemann solver of Roe [125] founded on the idea to linearise the nonlinear Riemann problem in an appropriate manner. Then, one can use the Riemann solvers of the Harten-Lax-Leer (HLL) family, [85], such as HLLC [68]. This family of RS is characterized by simplicity and great robustness. However, they smear out contact discontinuities and other internal waves, which is a big drawback for large PDEs systems. A significant advance for this problem has been made by Einfeldt *et al.*, who proposed the HLLEM Riemann solver [69], and by Toro *et al.*, who designed the HLLC method [149]. Both, HLLC and HLLEM are able to resolve steady contact discontinuities exactly, without compromising the robustness of the underlying HLL scheme. This property made HLLC one of the most widespread Riemann solvers nowadays, used in many computer codes. The strategy proposed by Osher and Solomon in [116] is based on path integrals and it is also well-suited for large PDE systems, since it is a so-called complete Riemann solver. However, the computational cost increases significantly. In addition, we also mention the local Lax-Friedrichs (LLF) or Rusanov method [126], which is a very simple but very robust approximate Riemann solver; it can be seen also as a particular case of the HLL RS. We conclude this list, mentioning that very recently the Osher and HLLEM Riemann solvers were extended also to non-conservative systems of hyperbolic PDEs, see [62, 65].

Additionally, the well-known CFL constraint on the time step  $\Delta t$  is necessary in order to ensure the stability of an explicit scheme, see [46]. In the one dimensional case, this stability condition reads:

$$\text{CFL} < 1, \quad \text{with} \quad \text{CFL} = |\lambda_{\max,i}| \frac{\Delta t}{\Delta x}, \quad (1.7)$$

where CFL is the so called Courant-Friedrichs-Lewy number and  $|\lambda_{\max,i}|$  is the absolute value of the maximum eigenvalue for every  $\mathbf{Q}_i^n$ .

Moreover, another essential property of any numerical method is the monotonicity. In general, if this crucial feature is not respected, the method produces non-physical oscillations in the presence of shocks and discontinuities. This phenomenon is unwanted, because it could generate a wrong physical understanding of the results or, even worse, the simulation could crash. In this sense, a barrier to the construction of monotonicity-preserving high order FV methods is given by the Godunov theorem [78], which states that *linear* monotone schemes can be at most first order accurate. However, over the years, researchers understood that this theorem can be circumvented by building *non-*

*linear* high order FV schemes, which respect the monotonicity property. Hence, a fundamental improvement in terms of accuracy has been achieved by van Leer [157], who proposed a *nonlinear* (data-dependent) reconstruction of piecewise polynomials of degree one in space and time, leading to the so-called MUSCL-Hancock method. The nonlinearity was achieved by a so-called slope limiter, and in the last decades, many of such slope limiters have been developed. The MUSCL-Hancock scheme belongs to the family of so-called *Total Variation Diminishing* (TVD) schemes, which are second-order accurate nonlinear methods for which in the scalar 1D case it can be proven that the total variation (TV) of the discrete solution does not increase in time.

Another significant progress is represented by the class of ENO schemes, proposed by Harten et al. in [86], and by the WENO scheme of Jiang and Shu [93]. ENO and WENO schemes are based on nonlinear data-dependent reconstruction, making use of several reconstruction stencils in order to find the least oscillatory reconstruction polynomial. High order of accuracy in space is achieved by extrapolating the values to the cell interfaces from the nonlinear reconstruction polynomial, while integration in time is typically done using high order TVD Runge-Kutta methods, see [40, 93].

Another successful strategy is the ADER scheme of Toro and Titarev [145, 146, 150, 151], where high order of accuracy in space and time are achieved in one step by using the approximate solution of the *generalized Riemann problem*.

The finite volume methods mentioned above can also be extended to nonlinear systems of hyperbolic conservation laws in multiple space dimensions, see [148]. We distinguish two different ways, depending on the subdivision of the computational domain. The first strategy is based on regular Cartesian grids and the other approach uses unstructured grids, typically adopted for the discretization of geometrically complex domains. For this last case we mention the possibility to discretize the system of PDEs also on moving conforming and non-conforming unstructured meshes, see e.g. [15, 16, 17, 76, 77].

To summarize, explicit finite volume methods are widely used for the numerical simulation of nonlinear hyperbolic systems of PDEs. They are characterized by great robustness and there are several strategies in order to obtain high order of accuracy in space and time. However, they are not suitable for the simulation of low Mach number or nearly incompressible flows, since the CFL condition (1.7) of an explicit FV scheme is based on the sound speed, which is contained in the eigenvalues of the compressible Euler system. Furthermore, in some particular situations, such as the simulation of compressible turbulent flows and the resulting acoustic waves, a data representation based merely on piecewise constant cell averages may not be enough. This makes piecewise polynomial data more desirable for the design of high order schemes.

### 1.1.2 Explicit Discontinuous Galerkin schemes

In the previous section, we explained that in the finite volume framework the computations are carried out using only cell averages and, consequently, high order finite volume schemes are based on reconstruction procedures of the piecewise constant values that lead to high order polynomials.

The discontinuous Galerkin schemes belong to the family of explicit discontinuous finite elements methods, and arbitrary high order of accuracy in space is easily obtained by working with piecewise polynomial basis functions. DG has first been proposed by Reed and Hill in 1973 for the investigation of neutron transport [124] and after fifteen years Cockburn and Shu recovered the method and they extended these schemes to general systems of nonlinear hyperbolic equations in a famous series of papers [41, 42, 43, 44, 44]. Considering the governing system of PDEs in eq. (1.1) and the computational domain  $\Omega$  introduced in the previous section, we multiply all the terms by a test function that consist in piecewise polynomials of degree  $P$ .  $\mathbf{Q}$  and  $\mathbf{F}$  are projected into a discrete space spanned by the same piecewise polynomial basis functions and within each spatial control volume they are expressed as follows:

$$\mathbf{Q}(x, t) = \sum_{l=0}^P \phi_l(x) \hat{\mathbf{Q}}_l(t), \quad \mathbf{F}(x, t) = \sum_{l=0}^P \phi_l(x) \hat{\mathbf{F}}_l(t) \quad (1.8)$$

where  $\hat{\mathbf{Q}}(t)$  and  $\hat{\mathbf{F}}(t)$  are the so-called degree of freedom of the vector of the conservative variables and of the flux vector, respectively. Then, we integrate over the control volume  $T_i$  and integration by parts produces

$$\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \phi \frac{\partial \mathbf{Q}}{\partial t} dx + \phi(x_{i+\frac{1}{2}}) \mathbf{F}_{i+\frac{1}{2}} - \phi(x_{i-\frac{1}{2}}) \mathbf{F}_{i-\frac{1}{2}} - \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \frac{\partial \phi}{\partial x} \mathbf{F}(\mathbf{Q}) dx = 0. \quad (1.9)$$

In eq. (1.9)  $\mathbf{F}_{i\pm\frac{1}{2}} = \mathbf{F}_{i\pm\frac{1}{2}}(\mathbf{Q}_{i\pm\frac{1}{2}}^-, \mathbf{Q}_{i\pm\frac{1}{2}}^+)$  are computed by using one of the numerical fluxes discussed in section 1.1.1 evaluated at the boundary extrapolated values  $\mathbf{Q}_{i\pm\frac{1}{2}}^\pm$  on the right and the left of the cell interface. The discretization of the PDE system continues using a TVD Runge-Kutta time integration. In most cases, the third order

accurate version of the TVD Runge-Kutta method is used. It reads as follows:

$$\begin{aligned}\hat{\mathbf{k}}_1 &= \hat{\mathbf{Q}}^n + \Delta t \mathbf{L}_h \left( \hat{\mathbf{Q}}^n \right), \\ \hat{\mathbf{k}}_2 &= \frac{3}{4} \hat{\mathbf{Q}}^n + \frac{1}{4} \hat{\mathbf{k}}_1 + \frac{1}{4} \Delta t \mathbf{L}_h \left( \hat{\mathbf{k}}_1 \right), \\ \hat{\mathbf{Q}}^{n+1} &= \frac{1}{3} \hat{\mathbf{Q}}^n + \frac{2}{3} \hat{\mathbf{k}}_2 + \frac{2}{3} \Delta t \mathbf{L}_h \left( \hat{\mathbf{k}}_2 \right),\end{aligned}\tag{1.10}$$

where the operator  $\mathbf{L}_h$  contains the space discretization

$$\mathbf{L}_h \left( \hat{\mathbf{Q}} \right)_i = -\frac{1}{\Delta x} \mathbf{M}^{-1} \left( \varphi(1) \mathbf{F}_{i+\frac{1}{2}} - \varphi(0) \mathbf{F}_{i-\frac{1}{2}} - \mathbf{K} \cdot \hat{\mathbf{F}}_i \right).\tag{1.11}$$

In addition, in eq. (1.11) the following universal matrices were introduced

$$M_{kl} = \int_0^1 \varphi(\xi) \varphi(\xi) d\xi, \quad K_{kl} = \int_0^1 \varphi'(\xi) \varphi(\xi) d\xi,\tag{1.12}$$

which can be simply precomputed. The success of the DG scheme is mainly based on the fact that it can easily obtain arbitrary high order of accuracy in space, and the third order of accuracy of the time integration is usually sufficient for many practical application. We also mention the so called ADER-DG method [53, 88, 122, 135], which achieves arbitrary high order of accuracy in space and time in a single step, which makes ADER-DG a fully discrete one-step scheme. However explicit DG schemes suffer of a very severe stability condition on the time step, because the Courant number depends on the polynomial degree  $P$  approximately as follows

$$\text{CFL} < \frac{1}{2P+1}.\tag{1.13}$$

Another important drawback is the fact that the DG scheme, as presented above, is *linear* in the sense of Godunov. Hence, spurious oscillations are inevitably generated in the vicinity of steep gradients and discontinuities. The design of appropriate limiters for high order DG methods is still an open subject of research and several different approaches have been proposed in the literature over the years. A first family of limiting strategies is the so called *a priori* limiting. It is divided into three main classes that are the *artificial viscosity method* [3, 114, 118], *(H)WENO limiting* [2, 94, 95] and *slope/moment limiters* [24, 49, 50, 100, 101, 108]. Recently, a novel family of *a posteriori* sub-cell finite volume limiting on uniform and adaptive Cartesian has been proposed in [64, 163]. This approach has been successfully extended also to fixed and moving unstructured grids in [14, 59]. Here, we also mention a family of DG schemes with *a priori* subcell finite volume limiting, see [111, 132, 133]. We furthermore provide a

non-exhaustive list of publications of explicit DG schemes in applied sciences considering the fields of computational fluid dynamics [7, 8, 9, 74], elastic wave propagation [57, 58, 143] and computational astrophysics [61, 73, 74, 164]. The major problem of explicit DG schemes is the time step restriction given by the stability condition in eq. (1.13). This issue can be circumvented or at least alleviated by using either *implicit* or *semi-implicit* time discretizations, which will be discussed in the next section 1.2.

## 1.2 Implicit and semi-implicit schemes

In this section, we discuss the families of implicit and semi-implicit schemes for the numerical approximation of PDEs. Instead of explicit schemes, they can also be applied to parabolic and elliptic PDEs and, in general, they are more efficient because they are either subject to some milder stability conditions (semi-implicit methods), or they are even unconditionally stable (fully implicit schemes).

Implicit methods are widespread in physical problems governed by linear or nonlinear parabolic equations, which describe, for example, the propagation of heat and pollutants, or groundwater flow in unsaturated porous media. We start by giving a very simple example of implicit method. For this purpose, we consider the one-dimensional diffusion equation

$$\frac{\partial C}{\partial t} + \frac{\partial F}{\partial x} = 0, \quad \text{with} \quad F = -D \frac{\partial C}{\partial x}, \quad (1.14)$$

where  $C$  is the concentration,  $D \geq 0$  is the diffusion coefficient and  $F$  is a flux term that depends on the  $D$  and the gradient of  $C$ . Introducing a computational domain  $\Omega = [x_L, x_R]$  subdivided into  $N_x$  equally spaced elements, a finite-difference approximation of eq. (1.14) with an implicit time discretization of the fluxes reads as follows

$$C_i^{n+1} + \frac{\Delta t}{\Delta x} (F_{i+\frac{1}{2}}^{n+1} - F_{i-\frac{1}{2}}^{n+1}) = C_i^n \quad \text{with} \quad F_{i+\frac{1}{2}}^{n+1} = -D_{i+\frac{1}{2}} \frac{C_{i+1}^{n+1} - C_i^{n+1}}{\Delta x}. \quad (1.15)$$

Moving unknowns at time level  $t^{n+1}$  to the left yields

$$-D_{i-\frac{1}{2}} \frac{\Delta t}{\Delta x^2} C_{i-1}^{n+1} + (1 + D_{i-\frac{1}{2}} \frac{\Delta t}{\Delta x^2} + D_{i+\frac{1}{2}} \frac{\Delta t}{\Delta x^2}) C_i^{n+1} - D_{i+\frac{1}{2}} \frac{\Delta t}{\Delta x^2} C_{i+1}^{n+1} = C_i^n. \quad (1.16)$$

After supplying proper boundary conditions, equation (1.16) repeated for all the indexes  $i = 1, \dots, N_x$  forms a three-diagonal linear algebraic system with  $N_x$  equations and  $N_x$  unknowns. Several strategies can be adopted in order to solve the linear systems that arise from the implicit discretization of PDEs. For example, *Gaussian elimination* and *LU factorization* methods that belong to the so called family of *direct methods*, see [123].

Here, we also mention the *Thomas algorithm*, see [123], which is a particular case of the Gaussian elimination method applied to tridiagonal systems of equations and which is particularly efficient in terms of memory and computational cost. Implicit time discretization can be easily extended to multiple space dimensions; however, the use of direct solvers for large linear algebraic systems is usually computationally very expensive. An efficient alternative is typically given by the class of *iterative linear solvers*. In particular, a robust and fast algorithm is the conjugate gradient (CG) method [87, 123], which can be applied when the system is symmetric and positive definite. If these two properties are not satisfied, one can for example use the GMRES method of Saad and Schultz [127].

The most important feature of fully implicit schemes is the fact that they are unconditionally stable and, consequently, they become very competitive in those situations where explicit schemes are restricted by very small time steps due to the CFL condition. Fully implicit methods can be adopted in order to solve numerically nonlinear PDE systems like the Navier-Stokes equations. In the context of implicit DG schemes, we refer the reader to the methods proposed by Bassi *et al.* [4, 5, 6, 110] and the fully implicit space time DG schemes of Van der Vegt *et al.* [98, 156]. Typically, solving nonlinear systems of PDEs using fully implicit schemes is computationally very expensive, since the implicit discretization yields strongly nonlinear algebraic systems with huge numbers of unknowns and with rather high condition number. In the context of high order DG schemes, the discretization is carried out mostly on *collocated grids*.

In order to keep using a large time step but increasing the computational efficiency, the semi-implicit approach is valid a strategy; hence a general nonlinear system of PDEs is *split* [152] as

$$\frac{\partial \mathbf{Q}}{\partial t} + \frac{\partial}{\partial x} \mathbf{F}_c(\mathbf{Q}) + \frac{\partial}{\partial x} \mathbf{F}_p(\mathbf{Q}) = 0, \quad (1.17)$$

where  $\mathbf{F}_c(\mathbf{Q})$  is the nonlinear *convective-type flux* and  $\mathbf{F}_p(\mathbf{Q})$  is a *pressure flux*. In the semi-implicit philosophy, the convective subsystem is updated *explicitly*, while the pressure system is integrated using an *implicit* time discretization. Hence, the time step is given by a *mild stability condition* based only on the fluid velocity and not based on the sound speed. In the framework of finite volume and finite difference methods, several very successful semi-implicit schemes have been developed by Casulli *et al.* for the shallow water and free-surface Navier-Stokes equations in a well-known series of papers [27, 28, 29, 33, 35, 36]. These methods are characterized by a staggered grid approach [84], where the free surface is located in the cell centers and the velocities are defined at the cell interfaces. Staggered semi-implicit finite volume methods have been recently also extended to the simulation of blood flow in systems of compliant arteries [34, 72, 141] and, successively, to the flow of compressible fluids in rigid and elastic pipes [56]. The idea of staggered meshes for semi-implicit finite volume schemes has

been extended also to the numerical simulation of the motion of fluids in porous media [19, 30, 31, 37], to the compressible Euler and Navier-Stokes equations [55] and to magnetohydrodynamics (MHD), see e.g. [66].

In the context of discontinuous Galerkin finite elements, the first high order semi-implicit DG schemes on *collocated meshes* were proposed by Dolejsi and Feistauer in [51, 52] for the Navier-Stokes equations. Further semi-implicit DG schemes on collocated grids have been proposed for the shallow water equations by Tumolo and Bonaventura in [154, 155].

A new class of semi-implicit DG methods on *staggered meshes* was first introduced by Dumbser and Casulli in [54] for the shallow water equations and this idea was successfully extended to unstructured meshes by Tavelli and Dumbser for different PDE systems in [136, 137, 138, 139, 140, 142]. High-order staggered semi-implicit space-time DG schemes for the solution of the incompressible Navier Stokes equations on uniform and adaptive Cartesian grids (AMR) have been proposed in [70, 71].

## 1.3 Targets of the thesis

The scope of the present work is to solve numerically partial differential equations using finite volume and high order finite element methods on staggered meshes with a semi-implicit time integration. Then, we want to use these new schemes in order to investigate some simple engineering applications.

Hence, at first we develop semi-implicit finite volume and finite elements schemes on staggered meshes in order to investigate pressurized fluids in rigid and elastic pipes. These numerical tools are very important in order to carry out simulations for biomedical, industrial and civil phenomena.

Another part of the work is devoted to the development of a novel shock capturing algorithm for the family of staggered semi-implicit high order Discontinuous Galerkin method. This tool is fundamental in order to perform successfully simulations characterized by strong shock waves which typically happen for shallow water flows or for acoustic phenomena at low Mach number regimes.

## 1.4 Structure of the thesis

The remaining chapters of this doctoral thesis are structured as follows:

- In Chapter 2 we discuss and compare explicit and semi-implicit FV methods for pressurized viscous compressible flows in elastic tubes with each other. Great

attention is dedicated to analyse the different types of frequency-dependent wall friction models available in literature. Moreover, we propose a new simple and computationally efficient approximation of Zielke's convolution integral to account for frequency-dependent friction.

- In Chapter 3 we still investigate compressible viscous pipe flow phenomena, and propose a novel staggered semi-implicit DG scheme for a cross-sectionally averaged 1D model and for a hydrostatic, axially symmetric  $2D_{xr}$  model. After the theoretical part, we perform several tests in order to validate the novel schemes.
- In Chapter 4 we deal with the 1D and 2D shallow water equations. First, we recall the basis of the staggered semi implicit DG method of Dumbser and Casulli [54]. Then, we proceed explaining the idea of a new *a posteriori* sub-cell finite volume limiter that is needed in order to deal with shock waves (bores) and the problem of wetting and drying in the shallow water equations. This is the first time ever that an *a posteriori* subcell finite volume limiter is used in the context of semi-implicit staggered DG schemes. In the last part of the chapter we carry out some benchmark simulations in order to validate the new family of methods.
- In Chapter 5 the new *a posteriori* subcell finite volume limiter is generalized to the 1D and 2D compressible Euler equations. We first start with the derivation of the unlimited SIDG schemes on staggered Cartesian meshes. Then we discuss the limited scheme and, finally, we execute the validation via several test cases that involve shock waves and other types of discontinuities.
- In Chapter 6 we draw the conclusions. In particular, we summarize the targets that have been achieved in this thesis and finally we discuss potential improvements and further steps that are left to future research.



## 2 A comparison of explicit and semi-implicit finite volume schemes for viscous compressible flows in elastic pipes in fast transient regime

In this chapter we compare the accuracy and computational efficiency of fully explicit and semi-implicit 1D and 2D finite volume schemes for the simulation of highly unsteady viscous compressible flows in laminar regime in axially symmetric compliant tubes. There are essentially two main classes of mathematical models that can be used to predict the pressure and velocity distribution along the tube: one class is based on the full compressible Navier-Stokes equations in an axially symmetric geometry, leading to a two-dimensional governing PDE system with moving boundaries, and the other class uses a simpler, cross-sectionally averaged version of the Navier-Stokes equations, which leads to a non-conservative PDE system in only one space dimension along the axial direction of the tube. Within the first class of models, the influence of the wall friction on the flow field is directly obtained from first principles, without any further modelling assumptions and is thus valid even for highly unsteady flows. In the second case, only averaged flow quantities are available, and it is well known from previous studies published in the literature that the correct representation of the wall friction needs to be *frequency dependent*, since the use of a simple steady friction model, like the classical Darcy-Weisbach law, is not sufficient to reproduce the wall friction effects in highly transient flows. For the cross-sectionally averaged Navier-Stokes equations, there are again two main classes of frequency-dependent wall friction models: convolution integral (CI) models and instantaneous acceleration (IA) models.

We provide a very thorough and critical comparison of all the above-mentioned methods for the simulation of highly oscillatory flows in rigid and compliant tubes concerning accuracy and computational efficiency. From our numerical results we can conclude that the convolution integral models are significantly superior to instantaneous acceleration models concerning accuracy. Furthermore, the CI models require only a slight computational overhead if they are properly implemented by solving a set of additional ODEs for appropriate auxiliary variables, instead of directly computing the convolution integrals. We also find that semi-implicit finite volume methods are clearly superior to conventional explicit finite volume schemes concerning computational efficiency, however, providing the same level of accuracy.

## 2.1 Mathematical models and numerical methods for pipe flow

### 2.1.1 Explicit path-conservative finite volume schemes for non-conservative hyperbolic systems

After having taken a cross-sectional average of the 3D compressible Navier-Stokes equations in an axially symmetric duct, the flow of a viscous and compressible fluid in a flexible pipe can be modeled by a simplified system of 1D partial differential equations (PDE) which is composed by the continuity and by the momentum equation. It is written as follows:

$$\begin{aligned} \frac{\partial}{\partial t}(\rho A) + \frac{\partial}{\partial x}(\rho AU) &= 0, \\ \frac{\partial}{\partial t}(\rho AU) + \frac{\partial}{\partial x}(\rho AU^2 + Ap) - p \frac{\partial A}{\partial x} &= -2\pi R\tau_w, \end{aligned} \quad (2.1)$$

where  $t \in \mathbb{R}_0^+$  is the time,  $x$  is the longitudinal coordinate which belongs to the 1D domain  $\Omega = [x_L, x_R] \subset \mathbb{R}$ ,  $\rho$  is the density,  $A$  is the cross sectional area of the tube,  $U$  is the mean velocity in the  $x$  direction,  $p$  is the mean pressure and  $\tau_w$  is the wall shear stress. This system is composed by two equations but it has five unknowns. Consequently three closure relations need to be introduced. The first one relates the pressure with density; here, we use a simple equation of state that is valid for a barotropic fluid:

$$p = p(\rho) = p_0 + c_0^2 (\rho - \rho_0), \quad (2.2)$$

where  $\rho_0$ ,  $p_0$  and  $c_0$  are the reference density, pressure and speed of sound of the fluid, respectively. The second closure relation connects the cross sectional area of the pipe with the pressure. Here we use a simple elastic model which is given by Hooke's law applied to a cross sectional area of a pipe [79, 105]:

$$A = \pi R^2, \quad R = R_0 \sqrt{1 + \frac{2Wp}{E_\infty}}. \quad (2.3)$$

In Eq. (2.3),  $R_0$  is the equilibrium radius,  $E_\infty$  is the Young's modulus of the material and  $W$  is a geometric parameter which contains the dependence on the wall thickness  $h_w$  and the Poisson ratio  $\nu_w$  of the tube, see [79, 105]:

$$W = 2 \frac{\left(\frac{h_w}{R_0}\right)^2 (1 + \nu_w) - (1 - 2\nu_w)}{\left(\frac{h_w}{R_0}\right)^2 - 1}. \quad (2.4)$$

In this work the material parameters  $h_w$  and  $\nu_w$  are assumed to be constant. Note that when  $E_\infty$  tends to infinity the tube becomes stiff and, consequently, the wall strain is negligible. The last closure regards the evaluation of  $\tau_w$ . In experiments characterized by fast transient flow the classical Darcy-Weisbach formula, which is only valid for the stationary regime, has to be supplemented by an additional term which contains the unsteady friction component for the computation of the tangential wall shear stress. Under the assumption of laminar flow the total wall shear stress is computed as  $\tau_w = \tau_s + \tau_u$ . Here,  $\tau_s$  is the quasi-steady friction term expressed as  $\tau_s = \rho\lambda|U|U/8$ , with  $\lambda = 64/Re$  being the usual Darcy-Weisbach friction factor for laminar flows given in terms of the Reynolds number  $Re$  and  $\tau_u$  is the unsteady wall shear stress. Its evaluation is discussed

in detail in section 2.1.3. Note that when  $\partial A/\partial x = 0$  and  $\tau_w = 0$  in (2.1), one obtains the well known 1D compressible Euler equations for barotropic fluids.

Introducing the vector of the physical variables defined as  $\mathbf{Q} = (\rho A, \rho AU)^T$ , system (2.1) can be written as follows:

$$\frac{\partial \mathbf{Q}}{\partial t} + \mathbf{J}(\mathbf{Q}) \frac{\partial \mathbf{Q}}{\partial x} + \mathbf{B}(\mathbf{Q}) \frac{\partial \mathbf{Q}}{\partial x} = \mathbf{S}(\mathbf{Q}). \quad (2.5)$$

where  $\mathbf{J}(\mathbf{Q}) = \partial \mathbf{f}(\mathbf{Q})/\partial \mathbf{Q}$  is the Jacobian matrix of the flux vector  $\mathbf{f} = (\rho AU, \rho AU^2 + Ap)^T = (Q_2, Q_2^2/Q_1 + Ap)^T$ .  $\mathbf{B}(\mathbf{Q})\partial \mathbf{Q}/\partial x$  represents the non-conservative term and the only non-zero element of the matrix  $\mathbf{B}$  is  $B_{21}$  which is equal to  $B_{21} = p/(\frac{c_0^2}{\partial A/\partial p} + \rho)$ ;  $\mathbf{S}(\mathbf{Q})$  is the source term and the non-zero term is given by the total wall shear stress that has to be properly written in conserved variables. It is possible to introduce the matrix  $\mathbf{A}(\mathbf{Q}) = \mathbf{J}(\mathbf{Q}) + \mathbf{B}(\mathbf{Q})$  and rewrite the system (2.5) as a quasi-linear system of PDEs as follows

$$\frac{\partial \mathbf{Q}}{\partial t} + \mathbf{A}(\mathbf{Q}) \frac{\partial \mathbf{Q}}{\partial x} = \mathbf{S}(\mathbf{Q}). \quad (2.6)$$

The diagonal matrix of eigenvalues  $\mathbf{\Lambda} = \text{diag}(\lambda_1, \lambda_2)$  and the matrix of right eigenvectors  $\mathbf{R}$  of the matrix  $\mathbf{A}$  read

$$\mathbf{\Lambda} = \begin{bmatrix} U - c' & 0 \\ 0 & U + c' \end{bmatrix} \quad \mathbf{R} = \begin{bmatrix} 1 & 1 \\ U - c' & U + c' \end{bmatrix} \quad \text{with} \quad c' = \frac{c_0}{\sqrt{1 + \frac{\rho c_0^2}{A} \frac{\partial A}{\partial p}}}, \quad (2.7)$$

where the parameter  $c'$  is the speed of sound modified by the effect of the wall elasticity and the derivative  $\partial A/\partial p$  can be directly derived from equation (2.3). Furthermore, it is important to recall that in the flux calculation the pressure has to be found from the known conservative quantities by solving the nonlinear scalar equation  $g(p) = Q_1 - \rho(p)A(p) = 0$ . In order to do this one can either use a Newton method or the more robust bisection method.

A grid composed by  $N_x$  equally spaced intervals  $T_i = [x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}]$  is introduced; the length of each element  $T_i$  is equal to  $\Delta x = x_{i+\frac{1}{2}} - x_{i-\frac{1}{2}} = (x_L - x_R)/N_x$ . The finite volume update based on a second order space-time TVD reconstruction for a non-conservative system of hyperbolic equations reads (see [65])

$$\mathbf{Q}_i^{n+1} = \mathbf{Q}_i^n - \frac{\Delta t}{\Delta x} (\mathbf{D}_{i+\frac{1}{2}}^- + \mathbf{D}_{i-\frac{1}{2}}^+) - \frac{\Delta t}{\Delta x} (\mathbf{f}_{i+\frac{1}{2}}^- - \mathbf{f}_{i-\frac{1}{2}}^+) - \frac{\Delta t}{\Delta x} \mathbf{B}(\mathbf{Q}_i^{n+\frac{1}{2}}) \Delta \mathbf{Q}_i^n + \Delta t \mathbf{S}(\mathbf{Q}_i^{n+\frac{1}{2}}), \quad (2.8)$$

where the  $\mathbf{D}_{i+\frac{1}{2}}^\pm$  are the so-called fluctuations and  $\mathbf{f}_{i+\frac{1}{2}}^{\pm, \pm}$  are the boundary extrapolated fluxes from within the element. The  $\Delta \mathbf{Q}_i^n$  are the slopes computed from the cell averages at time  $t^n$  by using the standard MINMOD slope limiter (applied componentwise) [148]

$$\Delta \mathbf{Q}_i^n = \text{minmod}(\mathbf{Q}_{i+1}^n - \mathbf{Q}_i^n, \mathbf{Q}_i^n - \mathbf{Q}_{i-1}^n). \quad (2.9)$$

Once that reconstructed values  $\mathbf{Q}_{i\pm\frac{1}{2}}^{n,\mp} = \mathbf{Q}_i^n \pm \Delta \mathbf{Q}_i^n$  at the time  $t^n$  are known, the quantities at the boundary and at the half time  $t^{n+\frac{1}{2}} = t^n + \frac{1}{2} \Delta t^n$  are given by the following time evolution:

$$\begin{aligned} \partial_t \mathbf{Q}_i^n &= - \frac{(\mathbf{f}(\mathbf{Q}_{i+\frac{1}{2}}^{n,-}) - \mathbf{f}(\mathbf{Q}_{i-\frac{1}{2}}^{n,+}))}{\Delta x} - \mathbf{B}(\mathbf{Q}_i^n) \frac{\Delta \mathbf{Q}_i^n}{\Delta x} + \mathbf{S}(\mathbf{Q}_i^n), \\ \mathbf{Q}_{i\pm\frac{1}{2}}^{n+\frac{1}{2},\mp} &= \mathbf{Q}_i^n \pm \frac{1}{2} \Delta \mathbf{Q}_i^n + \frac{\Delta t}{2} \partial_t \mathbf{Q}_i^n. \end{aligned} \quad (2.10)$$

We furthermore have used the abbreviation  $\mathbf{Q}_i^{n+\frac{1}{2}} = \mathbf{Q}_i^n + \frac{1}{2} \Delta t \partial_t \mathbf{Q}_i^n$ . Then, the fluctuations  $\mathbf{D}_{i+\frac{1}{2}}^\pm$  can be computed by using either the Osher Riemann solver [115, 116] or the HLL Riemann solver, both extended to non-conservative systems of hyperbolic equations using the path-conservative framework of Castro and Parés [26, 117]:

$$\mathbf{D}_{i+\frac{1}{2}}^\pm = \mathbf{D}_{\text{OSHER}}^\pm(\mathbf{Q}_{i+\frac{1}{2}}^{n+\frac{1}{2},-}, \mathbf{Q}_{i+\frac{1}{2}}^{n+\frac{1}{2},+}) \quad \text{or} \quad \mathbf{D}_{i+\frac{1}{2}}^\pm = \mathbf{D}_{\text{HLL}}^\pm(\mathbf{Q}_{i+\frac{1}{2}}^{n+\frac{1}{2},-}, \mathbf{Q}_{i+\frac{1}{2}}^{n+\frac{1}{2},+}). \quad (2.11)$$

We recall that for the path-conservative Osher scheme [67] the jump terms for two general states  $\mathbf{Q}_L$  and  $\mathbf{Q}_R$  read

$$\mathbf{D}_{\text{OSHER}}^\pm(\mathbf{Q}_L, \mathbf{Q}_R) = \frac{1}{2} \left( \int_0^1 \mathbf{A}(\psi(s)) ds \pm \int_0^1 |\mathbf{A}(\psi(s))| ds \right) (\mathbf{Q}_R - \mathbf{Q}_L), \quad (2.12)$$

when using the simple segment path  $\psi(\mathbf{Q}_L, \mathbf{Q}_R, s) = \mathbf{Q}_L + s(\mathbf{Q}_R - \mathbf{Q}_L)$ . However, the path integral appearing in the expression above can become quite cumbersome to evaluate for general PDE systems. With an abuse of notation, throughout the rest of the chapter we therefore define  $\mathbf{D}_{\text{OSHER}}^\pm(\mathbf{Q}_L, \mathbf{Q}_R)$  to be the *approximation* of (2.12) via *numerical quadrature*, i.e. as

$$\mathbf{D}_{\text{OSHER}}^\pm(\mathbf{Q}_L, \mathbf{Q}_R) = \frac{1}{2} \sum_j^G w_j (\mathbf{A}(\psi(s_j)) \pm |\mathbf{A}(\psi(s_j))|) (\mathbf{Q}_R - \mathbf{Q}_L), \quad (2.13)$$

where the path integral has been approximated using a Gauss-Legendre integration rule with nodes  $s_j$  and weights  $w_j$  on the unit interval  $[0, 1]$ . For systems composed of only two PDEs, as the one considered in (2.1), the HLLM Riemann solver proposed in [65] reduces to the HLL method, for which the jumps term (fluctuations) are given by the expressions

$$\begin{aligned} \mathbf{D}_{\text{HLL}}^-(\mathbf{Q}_L, \mathbf{Q}_R) &= - \frac{s_L}{s_R - s_L} [\mathbf{f}_R - \mathbf{f}_L + \mathbf{D}_*] + \frac{s_L s_R}{s_R - s_L} (\mathbf{Q}_R - \mathbf{Q}_L), \\ \mathbf{D}_{\text{HLL}}^+(\mathbf{Q}_L, \mathbf{Q}_R) &= + \frac{s_R}{s_R - s_L} [\mathbf{f}_R - \mathbf{f}_L + \mathbf{D}_*] - \frac{s_L s_R}{s_R - s_L} (\mathbf{Q}_R - \mathbf{Q}_L), \end{aligned} \quad (2.14)$$

with the abbreviation  $\mathbf{D}_* = \tilde{\mathbf{B}}(\mathbf{Q}_L, \mathbf{Q}_*)(\mathbf{Q}_* - \mathbf{Q}_L) + \tilde{\mathbf{B}}(\mathbf{Q}_*, \mathbf{Q}_R)(\mathbf{Q}_R - \mathbf{Q}_*)$  and where  $\mathbf{Q}_*$  is the intermediate HLL state computed as

$$\mathbf{Q}_* = \frac{1}{s_R - s_L} [(\mathbf{Q}_R s_R - \mathbf{Q}_L s_L) - (\mathbf{f}_R - \mathbf{f}_L) - \tilde{\mathbf{B}}(\mathbf{Q}_L, \mathbf{Q}_R)(\mathbf{Q}_R - \mathbf{Q}_L)], \quad (2.15)$$

and for  $\tilde{\mathbf{B}}(\mathbf{Q}_a, \mathbf{Q}_b)$  we have

$$\tilde{\mathbf{B}}(\mathbf{Q}_a, \mathbf{Q}_b) = \int_0^1 \mathbf{B}(\psi(\mathbf{Q}_a, \mathbf{Q}_b, s)) ds \approx \sum_j^G w_j \mathbf{B}(\psi(s_j))(\mathbf{Q}_a - \mathbf{Q}_b). \quad (2.16)$$

The wave speeds are estimated according to [65] as follows,

$$s_L = \min(0, \Lambda(\mathbf{Q}_L), \Lambda(\tilde{\mathbf{Q}})) \quad s_R = \max(0, \Lambda(\mathbf{Q}_R), \Lambda(\tilde{\mathbf{Q}})) \quad \text{with } \tilde{\mathbf{Q}} = \frac{1}{2}(\mathbf{Q}_R + \mathbf{Q}_L), \quad (2.17)$$

with  $\min(0, \Lambda, \tilde{\Lambda}) = \min(0, \lambda_1, \lambda_2, \tilde{\lambda}_1, \tilde{\lambda}_2)$  and  $\max(0, \Lambda, \tilde{\Lambda}) = \max(0, \lambda_1, \lambda_2, \tilde{\lambda}_1, \tilde{\lambda}_2)$ . As already mentioned above, in Eq. (2.13) and Eq. (2.16),  $s_j$  and  $w_j$  are the points and weights of an appropriate Gauss-Legendre quadrature formula. Here we consider  $G = 3$  and consequently  $s = (\frac{1}{2} - \frac{\sqrt{15}}{10}, \frac{1}{2}, \frac{1}{2} + \frac{\sqrt{15}}{10})$  and  $w = (\frac{5}{18}, \frac{8}{18}, \frac{5}{18})$ . For all the details see [67]. The scheme obeys the stability condition  $\Delta t = \text{CFL} \frac{\Delta x}{\max|\Lambda(\mathbf{Q}_i^n)|}$  with  $\text{CFL} \leq 1$  and it is second order accurate both in space and in time. This stability condition is based on the speed of sound of the flow, but for incompressible flows, the ratio of flow speed and sound speed (the Mach number) tends to zero. This means that in the incompressible limit explicit time stepping schemes become increasingly inefficient and a different strategy, such as the semi-implicit approach explained in [55, 56], has to be introduced in order to perform faster simulations.

Furthermore, as shown in [65, 104], the path-conservative HLLEM scheme and the path-conservative Osher Riemann solver are well balanced for certain classes of non-conservative PDE systems. In particular, these schemes exactly preserve steady state solutions with zero velocity and constant pressure in ducts with variable reference cross section ( $A_0 = \pi R_0^2$ ) in longitudinal direction. However, for the sake of simplicity we only consider the case  $R_0 = \text{const}$ , since the main focus of this chapter is on the detailed study of the influence of the wall friction in highly transient flows.

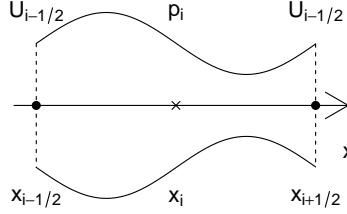
### 2.1.2 Semi-implicit finite volume schemes

In order to increase the computational efficiency of the simulation, especially for low Mach number flows, an alternative semi-implicit discretization can be used in order to approximate the PDE system (2.1). Here, for convenience, it is rewritten as follows:

$$\frac{\partial}{\partial t}(\rho A) + \frac{\partial}{\partial x}(\rho A U) = 0, \quad (2.18a)$$

$$\frac{\partial}{\partial t}(\rho A U) + \frac{\partial}{\partial x}(\rho A U^2) + A \frac{\partial p}{\partial x} = -2\pi R \tau_s - 2\pi R \tau_u. \quad (2.18b)$$

First, a *staggered grid* is introduced in order to discretize the one-dimensional domain. Hence, the pipe of length  $L$  is divided into  $N_x$  intervals of constant length  $\Delta x = L/N_x$ . The pressure is located at the cell centers  $x_i$ , while the velocities and the mass fluxes are defined at the interfaces  $x_{i \pm \frac{1}{2}}$  of the control volumes (see Fig. 2.1). In addition, the  $\theta$ -method is introduced for the time discretization, where  $\theta$  is an implicitness parameter taken in the range  $0.5 \leq \theta \leq 1$  for stability. Moreover, when  $\theta = 1$  the scheme is of first order of accuracy in time, while when  $\theta = 0.5$  one has a Crank-Nicolson type scheme of second order. For example, the  $\theta$ -method applied to the pressure gives  $p_i^{n+\theta} = \theta p_i^{n+1} + (1 - \theta) p_i^n$ . The  $\theta$ -method is widely used in semi-implicit schemes, see [34, 55, 56, 72, 136, 141], because it is quite a simple and cheap strategy to achieve second order



**Figure 2.1:** Grid for the semi-implicit 1D scheme: according to the staggered approach, the pressure is defined in the cell barycenter while the velocity is defined in the cell edges

of accuracy in time.

The semi-implicit discretization for the continuity equation yields [56]

$$M(p_i^{n+1}) = M(p_i^n) - \Delta t(Q_{i+\frac{1}{2}}^{n+\theta} - Q_{i-\frac{1}{2}}^{n+\theta}), \quad (2.19)$$

where  $M(p_i) = \rho(p_i)A(p_i)\Delta x$  is the mass of fluid contained in the  $i$ -th cell and  $Q_{i+\frac{1}{2}}^{n+1} = \rho_{i+\frac{1}{2}}^n A_{i+\frac{1}{2}}^n U_{i+\frac{1}{2}}^{n+1}$  is the mass flow rate with  $\rho_{i+\frac{1}{2}}^n = \max[0, \frac{1}{2}\rho(p_i^n) + \frac{1}{2}\rho(p_{i+1}^n)]$  and  $A_{i+\frac{1}{2}}^n = \max[0, \frac{1}{2}A(p_i^n) + \frac{1}{2}A(p_{i+1}^n)]$ . At the same time, the discrete momentum equation reads

$$\begin{aligned} \frac{Q_{i+\frac{1}{2}}^{n+1} - FQ_{i+\frac{1}{2}}^n}{\Delta t} &= -A_{i+\frac{1}{2}}^n \frac{\Delta t}{\Delta x} (p_{i+1}^{n+\theta} \\ &\quad - p_i^{n+\theta}) - 2\pi R_{i+\frac{1}{2}}^n \rho_{i+\frac{1}{2}}^n \lambda_{i+\frac{1}{2}}^n \frac{U_{i+\frac{1}{2}}^n U_{i+\frac{1}{2}}^{n+1}}{8} - 2\pi R_{i+\frac{1}{2}}^n (\tau_u^n)_{i+\frac{1}{2}}, \end{aligned} \quad (2.20)$$

where  $FQ_{i+\frac{1}{2}}^n$  is an explicit and nonlinear operator for the convective terms. It can be computed in many ways and a simple choice can be a finite volume method based on the Rusanov flux, as done in [56]:

$$\begin{aligned} FQ_{i+\frac{1}{2}}^n &= Q_{i+\frac{1}{2}}^n - \frac{\Delta t}{\Delta x} (f_{i+1}^n - f_i^n), \quad \text{where} \\ f_i &= \frac{1}{2} (U_{i+\frac{1}{2}}^n Q_{i+\frac{1}{2}}^n + U_{i-\frac{1}{2}}^n Q_{i-\frac{1}{2}}^n) - \frac{1}{2} S_{\max} (Q_{i+\frac{1}{2}}^n - Q_{i-\frac{1}{2}}^n), \\ S_{\max} &= 2 \max(|U_{i-\frac{1}{2}}^n|, |U_{i+\frac{1}{2}}^n|). \end{aligned} \quad (2.21)$$

The evaluation of term  $\tau_u^n$  is explained in section 2.1.3. Eq. (2.20) can be rewritten as

$$Q_{i+\frac{1}{2}}^{n+1} = G_{i+\frac{1}{2}}^n - \theta A_{i+\frac{1}{2}}^n \frac{\Delta t}{\Delta x} (p_{i+1}^{n+1} - p_i^{n+1}) - \Delta t \gamma_{i+\frac{1}{2}}^n Q_{i+\frac{1}{2}}^{n+1}, \quad (2.22)$$

where the term  $\gamma_{i+\frac{1}{2}}^n = \frac{2\pi R_{i+\frac{1}{2}}^n \lambda_{i+\frac{1}{2}}^n |U_{i+\frac{1}{2}}^n|}{8A_{i+\frac{1}{2}}^n} \geq 0$  takes into account the effects of quasi-steady friction using the Darcy-Weisbach law. Here  $\lambda_{i+\frac{1}{2}}^n = 64/Re$  denotes the (local) Darcy friction factor in terms of the Reynolds number of the flow. Moving  $Q_{i+\frac{1}{2}}^{n+1}$  to the left hand side, the previous expression becomes

$$Q_{i+\frac{1}{2}}^{n+1} = \left( \frac{G}{1 + \Delta t \gamma} \right)_{i+\frac{1}{2}}^n - \theta \frac{\Delta t}{\Delta x} \left( \frac{A}{1 + \Delta t \gamma} \right)_{i+\frac{1}{2}}^n (p_{i+1}^{n+1} - p_i^{n+1}), \quad (2.23)$$

where the quantity  $G_{i+\frac{1}{2}}^n$  includes the nonlinear convection and the known terms at time  $t^n$ :

$$G_{i+\frac{1}{2}}^n = FQ_{i+\frac{1}{2}}^n - (1 - \theta) A_{i+\frac{1}{2}}^n \frac{\Delta t}{\Delta x} (p_{i+1}^n - p_i^n) - 2\pi R_{i+\frac{1}{2}}^n (\tau_u)_{i+\frac{1}{2}}^n \Delta t. \quad (2.24)$$

Substitution of Eq. (2.23) into Eq. (2.19) yields:

$$M(p_i^{n+1}) - \theta^2 \frac{\Delta t^2}{\Delta x} \left[ (p_{i+1}^{n+1} - p_i^{n+1}) \left( \frac{A}{1 + \Delta t \gamma} \right)_{i+\frac{1}{2}}^n - (p_i^{n+1} - p_{i-1}^{n+1}) \left( \frac{A}{1 + \Delta t \gamma} \right)_{i-\frac{1}{2}}^n \right] = b_i^n, \quad (2.25)$$

which can be written more compactly using the following matrix vector notation

$$\mathbf{M}(\mathbf{p}^{n+1}) + \mathbf{T}\mathbf{p}^{n+1} = \mathbf{b}(\mathbf{p}^n), \quad (2.26)$$

where  $\mathbf{M}$  is a vector function that contains the non-linearity,  $\mathbf{T}$  is a symmetric three-diagonal matrix,  $\mathbf{p}^{n+1}$  is the vector of the unknown pressures and  $\mathbf{b}^n$  is the known right hand side term. System (2.26) is efficiently solved by using either the Newton algorithm of Brugnano and Casulli [20, 21], or the more general nested Newton algorithm of Casulli and Zanolli [37, 38, 56], if more complex equations of state are needed, see [55, 56]. When the pressure is known, the numerical mass flux can be easily computed by using Eq. (2.23) and, consequently, the new velocity is equal to  $U_{i+\frac{1}{2}}^{n+1} = Q_{i+\frac{1}{2}}^{n+1} / (\rho_{i+\frac{1}{2}}^n A_{i+\frac{1}{2}}^n)$ . Moreover, the density and the cross sectional area at the new time are updated using the closures in Eq. (2.2) and in Eq. (2.3).

To improve the information that regards the radial velocity profile, one can use a semi-implicit scheme for a two-dimensional model such as the one suggested in [56]. This method is an extension

## 2 Explicit and semi-implicit FV schemes for compressible flows in elastic pipes

to weakly compressible flows of the family of semi-implicit methods for blood flow presented in a series of recent papers [34, 72, 141]. Under the hypothesis of hydrostatic radial pressure equilibrium and when the longitudinal scale is much larger than the radial one, the motion of a compressible barotropic fluid in a circular elastic duct is described by the following system of equations, which is a simplification of the Navier-Stokes equation in cylindrical coordinates:

$$\begin{aligned} \frac{\partial \rho r}{\partial t} + \frac{\partial}{\partial x}(\rho u r) + \frac{\partial}{\partial r}(\rho r w) &= 0, \\ \rho \left( \frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} + w \frac{\partial u}{\partial r} \right) + \frac{\partial p}{\partial x} &= \frac{\mu}{r} \frac{\partial}{\partial r} \left( r \frac{\partial u}{\partial r} \right), \end{aligned} \quad (2.27)$$

where  $r$  is the radial coordinate,  $\mu$  is the dynamic viscosity, and  $u$  and  $w$  are the velocity components in  $x$  and  $r$  direction, respectively.

The kinematic boundary condition at the moving wall yields (see [34, 72, 141])

$$\frac{\partial R}{\partial t} + u \frac{\partial R}{\partial x} - w = 0 \quad \text{for } r = R. \quad (2.28)$$

The first equation of system (2.27) is integrated over the cross section and using the kinematic boundary condition the following relation is obtained:

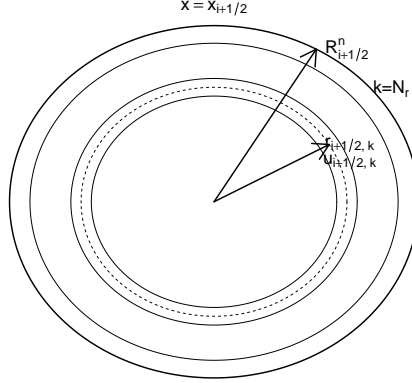
$$\frac{\partial \rho A}{\partial t} + 2\pi \frac{\partial}{\partial x} \int_0^R \rho u r dr = 0. \quad (2.29)$$

The quantities  $A$  and  $\rho$  can be computed by using the same closures introduced in the previous section. Similarly to the previous method, in axial direction there are  $N_x$  cells of length  $\Delta x$  and the pressure  $p_i$  is located at the center of these control volumes in the position  $x_i$ . The radial grid is composed of ring elements of size  $\Delta r_{i+\frac{1}{2},k}^n = r_{i+\frac{1}{2},k+\frac{1}{2}}^n - r_{i+\frac{1}{2},k-\frac{1}{2}}^n$  with  $k = 1, 2, \dots, K_{i+\frac{1}{2}}^n$  where  $K_{i+\frac{1}{2}}^n$  is the number of active rings at the cell interface  $x_{i+\frac{1}{2}}$ .  $K_{i+\frac{1}{2}}^n$  is bounded between 1 and  $N_r$ , which is the maximum number of ring elements. For simplicity, we will always use the maximum number of ring elements ( $K_{i+\frac{1}{2}} = N_r$ ). However we want to underline that in principle the semi-implicit 2D scheme allows to modify the number of active layers according to the wall deformation. The axial velocity  $u$  is located at the edge  $x_{i+\frac{1}{2}}$  and at the midpoint  $r_{i+\frac{1}{2},k}$  while  $w_{i,k+\frac{1}{2}}$  is defined in the cell barycenter  $x_i$  and at the edge of the radial segment  $r_{k+\frac{1}{2}}$  (see figures 2.2 and 2.3). This discretization of the computational domain yields again a staggered grid. The momentum equation is discretized using a semi-implicit approach [56]:

$$\begin{aligned} \frac{q_{i+\frac{1}{2},k}^{n+1} - F q_{i+\frac{1}{2},k}^n}{\Delta t} &= -a_{i+\frac{1}{2},k}^n \frac{p_{i+1}^{n+\theta} - p_i^{n+\theta}}{\Delta x}, \\ &+ 2\pi\mu \left( r_{i+\frac{1}{2},k+\frac{1}{2}}^n \frac{u_{i+\frac{1}{2},k+1}^{n+\theta} - u_{i+\frac{1}{2},k}^{n+\theta}}{\Delta r_{i+\frac{1}{2},k+\frac{1}{2}}^n} - r_{i+\frac{1}{2},k-\frac{1}{2}}^n \frac{u_{i+\frac{1}{2},k}^{n+\theta} - u_{i+\frac{1}{2},k-1}^{n+\theta}}{\Delta r_{i+\frac{1}{2},k-\frac{1}{2}}^n} \right), \end{aligned} \quad (2.30)$$

where  $q_{i+\frac{1}{2},k}^{n+1} = \rho_{i+\frac{1}{2}}^n a_{i+\frac{1}{2}}^n u_{i+\frac{1}{2}}^{n+1}$  and  $a_{i+\frac{1}{2},k}^n = 2\pi r_{i+\frac{1}{2},k}^n \Delta r_{i+\frac{1}{2}}^n$  are respectively the mass flow and the area of the  $k$ -th ring in position  $x_{i+\frac{1}{2}}$ .  $F q_{i+\frac{1}{2},k}^n$  is an explicit and non-linear operator





**Figure 2.2:** Radial grid for the axial velocity  $u$

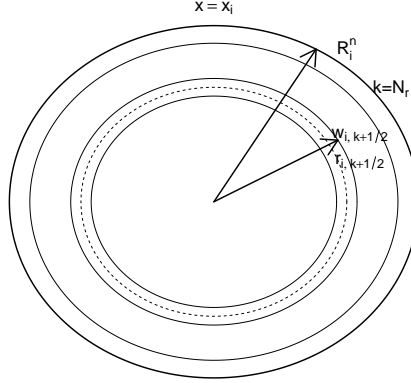
for the convective terms. The discrete momentum equation can be rewritten in a more compact form obtaining the vector of the mass fluxes [56]:

$$\mathbf{Q}_{i+\frac{1}{2}}^{n+1} = [\mathbf{D}^{-1} \mathbf{G}]_{i+\frac{1}{2}}^n - \theta \frac{\Delta t}{\Delta x} (p_{i+1}^{n+1} - p_i^{n+1}) [\mathbf{D}^{-1} \mathbf{A}]_{i+\frac{1}{2}}^n, \quad (2.31)$$

where  $\mathbf{Q}_{i+\frac{1}{2}}^{n+1} = (q_{i+\frac{1}{2},1}, q_{i+\frac{1}{2},2}, \dots, q_{i+\frac{1}{2},K_{i+\frac{1}{2}}^n})^T$  and  $\mathbf{A}_{i+\frac{1}{2}}^{n+1} = (a_{i+\frac{1}{2},1}, \dots, a_{i+\frac{1}{2},K_{i+\frac{1}{2}}^n})^T$  are column vectors of the mass and of the area of the rings.  $\mathbf{D}_{i+\frac{1}{2}}^{n+1}$  is a three-diagonal matrix that accounts for the coefficients of the viscous terms at the time  $t^n + \Delta t$  (see [34, 56, 141]). The column vector  $\mathbf{G}_{i+\frac{1}{2}}^n$  contains the non linear convective terms and the known quantities at time  $t^n$ . Eq. (2.29) is approximated by using a finite volume discretization. Also here the  $\theta$ -method is used in order to improve time accuracy,

$$M(p_i^{n+1}) = M(p_i^n) - \Delta t (\mathbf{1}^T \cdot \mathbf{Q}_{i+\frac{1}{2}}^{n+\theta} - \mathbf{1}^T \cdot \mathbf{Q}_{i-\frac{1}{2}}^{n+\theta}), \quad (2.32)$$

where we have used the notation  $\mathbf{1}^T = (1, 1, \dots, 1)$  for a row vector consisting only of ones. Substituting the discrete momentum equation (2.31) into the discrete continuity equation Eq. (2.32),



**Figure 2.3:** Radial grid for the radial velocity  $w$

one obtains the following mildly non-linear system:

$$M(p_i^{n+1}) - \theta^2 \frac{\Delta t^2}{\Delta x} \{(p_{i+1}^{n+1} - p_i^{n+1})[\mathbf{1}^T \mathbf{D}^{-1} \mathbf{A}]_{i+\frac{1}{2}}^n - (p_i^{n+1} - p_{i-1}^{n+1})[\mathbf{1}^T \mathbf{D}^{-1} \mathbf{A}]_{i-\frac{1}{2}}^n\} = b_i^n, \quad (2.33)$$

which has the same form of system (2.26) and is therefore solved in the same way at the aid of the (nested) Newton method. Once the pressure is known, the axial velocity is computed with Eq. (2.31), while the radial velocity is obtained from the discretization of the continuity equation (2.1). For more details about the numerical method, see [34, 56, 72, 141].

The stability condition of the semi-implicit schemes corresponds to the stability condition of the method used for the calculation of nonlinear convective terms. In general, the time step restriction is only based on the speed of the flow and *not* on the sound speed. Consequently, these schemes are very efficient, especially in the case of low Mach number flows. Moreover, if the convective terms can be neglected, one can simply set  $Fq = q$ , and in this case the schemes are *unconditionally stable*. However, the time step  $\Delta t$  should be chosen small enough to reduce the numerical viscosity and to resolve unsteady flow features properly.

In addition, it is rather simple to show that the semi-implicit schemes satisfy the so-called C-property, i.e. that these schemes are well-balanced. Consider the case where the cross sectional reference area of the pipe  $A_0 = \pi R_0^2$  is not constant in longitudinal direction, while the pressure is

constant ( $p_i^n = \text{const.}$ ) and the velocity is zero everywhere ( $U_{i+\frac{1}{2}}^n = 0$ ). In this case the discrete pressure gradients in (2.20) are zero and also the non-linear convective terms are zero, which can be easily seen by inserting  $U_{i+\frac{1}{2}}^n = 0$  into Eq. (2.21). Furthermore, also the unsteady friction terms vanish. As a result, the discrete mass fluxes  $Q_{i+\frac{1}{2}}^{n+1} = 0$  at time  $t^{n+1}$  remain zero and thus the pressure remains unchanged ( $p_i^{n+1} = p_i^n = \text{const.}$ ), i.e. the steady state solution is exactly preserved.

### 2.1.3 Unsteady friction models

#### 2.1.3.1 Convolution integral methods (CI)

The analytic expression for the calculation of the unsteady loss is given by the convolution integral of Zielke, see [162],

$$\tau_u(x, t) = \frac{2\mu}{R} \int_0^t \mathcal{W}(t - t^*) \frac{\partial U(x, t^*)}{\partial t} dt^*, \quad (2.34)$$

which is also called *frequency dependent friction*, since Zielke derived this expression in the complex domain after applying the Laplace transform. Moreover, he assumed incompressible and laminar flow in a rigid pipe and implemented his formula with the method of characteristics. In Eq. (2.34) the acceleration is multiplied by a weighting function  $\mathcal{W}$  of the following form

$$\mathcal{W}(\hat{t}) = \begin{cases} \sum_{j=1}^{j=6} m_j \hat{t}^{0.5(j-2)} & \text{for } \hat{t} \leq 0.02, \\ \sum_{j=1}^{j=5} e^{-n_j \hat{t}} & \text{for } \hat{t} \geq 0.02, \end{cases} \quad (2.35)$$

where the coefficients are given by  $(m_1, \dots, m_6) = (0.282095, -1.250000, 1.057855, 0.937500, 0.396696, -0.351563)$  and  $(n_1, \dots, n_5) = (26.3744, 70.8493, 135.0198, 218.9216, 322.5544)$ , while  $\hat{t} = \nu t / R^2$  is the dimensionless time. Note that the domain of definition of the weighting function is  $0 < \hat{t} < \infty$ . The numerical evaluation of the convolution integral needs a very large memory amount and makes this approach computationally very expensive. Consequently, another more efficient approach is necessary. First, Trikha proposed in [153] an approximated weighting function of the following form:

$$\mathcal{W}_{app}(\hat{t}) = \sum_{j=1}^{N_w} m_j e^{-n_j \hat{t}}, \quad (2.36)$$

with  $N_w = 3$ ,  $(m_1, m_2, m_3) = (40.0, 8.1, 1)$  and  $(n_1, n_2, n_3) = (8000, 200, 26.4)$ . Using a weighting function written as series of exponential functions, the unsteady wall shear is calculated with the formula of Trikha [153]

$$\tau_u^n \approx \frac{2\mu}{R} \sum_{j=1}^3 y_j^n = \frac{2\mu}{R} \left( \sum_{j=1}^3 e^{-n_j \frac{\nu \Delta t}{R^2}} y_j^{n-1} + \sum_{j=1}^3 m_j [U^n - U^{n-1}] \right), \quad (2.37)$$

which improved the efficiency but not the accuracy because the weighting function of Trikha does not reproduce well the original function. Successively Kagawa et al. [96] proposed a more accurate

## 2 Explicit and semi-implicit FV schemes for compressible flows in elastic pipes

weighting function with  $N_w = 10$  terms (valid for  $6.31 \cdot 10^{-6} \leq \hat{t} < \infty$ ) and another formula to approximate the Zielke convolution integral:

$$\tau_u^n \approx \frac{2\mu}{R} \sum_{j=1}^{N_w} y_j^n = \frac{2\mu}{R} \left( \sum_{j=1}^{N_w} e^{-n_j \frac{\nu \Delta t}{R^2}} y_j^{n-1} + \sum_{j=1}^{N_w} m_j e^{-n_j \frac{\nu \Delta t}{2R^2}} [U^n - U^{n-1}] \right). \quad (2.38)$$

The coefficients of the Kagawa weighting function are  $(n_1, \dots, n_{10}) = (26.3744, 72.8033, 187.424, 536.626, 1570.60, 4618.13, 13601.1, 40082.5, 118153, 348316)$  and  $(m_1, \dots, m_{10}) = (1, 1.16725, 2.20064, 3.92861, 6.78788, 11.6761, 20.0612, 34.4541, 59.4541, 101.59)$ .

Other authors proposed other values for the coefficients  $m_j$  and  $n_j$  for the approximated weighting function Eq. (3.13), such as Vitkovsky et al. [160] and Vardy and Brown [159]. In addition, Urbanowicz and Zarzycky [165] recently proposed a very accurate function with  $N_w = 26$ , defined in the range  $10^{-9} \leq \hat{t} < \infty$ :  $(n_1, \dots, n_{26}) = (26.3744, 70.8493, 135.0198, 218.9216, 322.5544, 499.148, 1072.543, 2663.013, 6566.001, 15410.459, 35414.779, 80188.189, 177078.960, 388697.936, 850530.325, 1835847.582, 3977177.832, 8721494.927, 19120835.527, 42098544.588, 92940512.285, 203458923, 445270063.893, 985067938, 2166385707.058, 4766167206.672)$  and  $(m_1, \dots, m_{26}) = (1, 1, 1, 1, 1, 2.141, 4.544, 7.566, 11.299, 16.531, 24.794, 36.229, 52.576, 78.150, 113.873, 165.353, 247.915, 369.561, 546.456, 818.871, 1209.771, 1770.756, 2651.257, 3968.686, 5789.566, 8949.468)$ . For a very nice review of convolution integral theory and of its approximation see [165].

### 2.1.3.2 Instantaneous acceleration methods (IA)

Another class of unsteady friction models is the one called Instantaneous Acceleration based models and is based on the hypothesis that the unsteady wall shear stress is directly proportional to the acceleration of the flow. To the knowledge of the authors, the most complete model of this family is the one of Brunone et al. which is formulated as follows [22, 23, 120]:

$$\tau_{us} = \frac{\rho D K_{Bru}}{4} \left( \frac{\partial U}{\partial t} + \text{sign}(U) \frac{\partial U}{\partial x} c_0 \frac{\partial U}{\partial x} \right). \quad (2.39)$$

Several values have been proposed for the coefficient  $K_{Bru}$  and here we use the expression suggested by Vardy and Brown in [158] because it seems to be the only one related to the Reynolds number

$$K_{Bru} = 2 \sqrt{\frac{7.41}{Re^\chi}}, \quad \chi = \log\left(\frac{14.3}{Re^{0.05}}\right). \quad (2.40)$$

### 2.1.4 A new efficient approximation of the Zielke convolution integral

The numerical integration of a convolution integral can be very computationally expensive, especially when long time simulations are required. Here, we derive a new simplification of the model of Zielke that leads to a new efficient formula for the evaluation of the unsteady wall shear stress. Our idea follows the approach used in the work of Moczo and Kristek that simplified convolution integrals in order to compute efficiently material-independent anelastic functions for seismic wave propagation [112]. Substituting a weighting function expressed as a sum of exponential functions

(eq. (3.13)) into the convolution integral, the total wall shear is computed as  $\tau_u = \sum_j^{N_w} \tau_j$ , where the  $j$ -th component is written as follows

$$\tau_j(t) = \frac{2\mu}{R} \int_0^t m_j e^{-\frac{n_j \nu}{R^2}(t-t^*)} \frac{dU(t^*)}{dt} dt^*. \quad (2.41)$$

The coefficients  $m_j$  and  $n_j$  can be taken from any of the approximated weighting functions previously cited. Deriving with respect to time the left and the right side of the previous equation and applying the Leibniz rule, we obtain the expression

$$\frac{d}{dt} \tau_j(t) = -\frac{2\mu}{R} \int_0^t \frac{dU(t^*)}{dt} m_j \frac{n_j \nu}{R^2} e^{-\frac{n_j \nu}{R^2}(t-t^*)} dt^* + \frac{2\mu}{R} \frac{dU(t)}{dt} m_j \quad (2.42)$$

that can be formally rewritten as an ordinary differential equation

$$\frac{d}{dt} \tau_j(t) = -\frac{n_j \nu}{R^2} \tau_j + \frac{2\mu}{R} \frac{dU(t)}{dt} m_j. \quad (2.43)$$

Since the coefficients  $n_j$  can assume very large values [96, 159, 160, 165], the source terms in the ODE (2.43) are stiff. Hence, for efficiency and stability reasons, the ODE (2.43) is approximated by using the *implicit* Euler method

$$\frac{\tau_j^n - \tau_j^{n-1}}{\Delta t} = -\frac{n_j \nu}{R^2} \tau_j^n + \frac{2\mu}{R} \frac{U^n - U^{n-1}}{\Delta t} m_j, \quad (2.44)$$

and finally, after solving for the quantity  $\tau_j^n$ , the unsteady friction term at time  $t^n$  reads:

$$\tau_u^n = \sum_{j=1}^{N_w} \tau_j^n = \sum_{j=1}^{N_w} \frac{\tau_j^{n-1} + \frac{2\mu}{R} m_j (U^n - U^{n-1})}{1 + \frac{n_j \nu}{R^2} \Delta t}. \quad (2.45)$$

Here, we deliberately omit the spatial index for the averaged velocity  $U$ , since the above formula (3.12) applies to both, the path-conservative finite volume scheme on collocated grids, as well as to the semi-implicit finite volume method on staggered grids.

## 2.2 Numerical results

In this section we show and compare the computational results that have been obtained by applying the previously presented numerical schemes. We use some abbreviations: PC1D for the explicit path conservative scheme applied to the 1D model, SI1D and SI2D for the semi-implicit schemes applied to the one-dimensional and to the two dimensional PDE, respectively. The unsteady friction model of Urbanowicz and Zarzycki is abbreviated by UZ model. Furthermore, Kagawa ODE and UZ ODE indicate that the new approximation of the Zielke convolution integral based on the ODE solver (3.12) has been used, and that it has been implemented by using the coefficients of the weighting function of Kagawa et al. [96] and of Urbanowicz and Zarzycky [165], respectively. In addition, the Osher Riemann solver and the HLL Riemann solver give very similar results for the 1D PDE system, which can be explained by the fact that the system does not contain any linearly degenerate intermediate field. Consequently, in the next plots the numerical solution for the explicit path-conservative finite volume scheme is referred to the one computed by using the Osher Riemann solver, in order to keep the images clear and readable. The differences between the methods can be appreciated in terms of accuracy and computational time for the simulations.

### 2.2.1 Womersley test

For the first test considered the fluid is assumed incompressible ( $c_0 = 10^{20}$ ) and the pipe rigid ( $E_\infty = 10^{20}$ ). Consequently, this test case can be numerically performed only by using the semi-implicit schemes because when the speed of sound  $c_0$  and thus also the maximum eigenvalue ( $\max|\Lambda(\mathbf{Q}_1^n)|$ ) tends to infinity, the system (2.1) is no longer hyperbolic.

An oscillating pressure gradient of the following form

$$p_{in}(t) = \Re(p_{out} + LA_w e^{j\omega t}) = p_{out} + LA_w \cos(\omega t) \quad (2.46)$$

is imposed in a pipe of length  $L$  and radius  $R$ . The pressure wave is characterized by an amplitude  $A_w$  and by an angular frequency  $\omega = 2\pi f$ . Due to the incompressibility and rigidity conditions, the pressure varies linearly along the longitudinal coordinates. The axial velocity  $u$  is constant in longitudinal direction while the radial profile is given by the analytic solution of Womersley [161]

$$u(r, t) = \Re\{A_w \frac{1}{j\omega} [1 - \frac{J_o(\hat{r}j^{3/2}Wo)}{J_o(j^{3/2}Wo)]} e^{j\omega t}\}, \quad (2.47)$$

where  $j = \sqrt{-1}$  is the imaginary unit,  $Wo = R\sqrt{\frac{\omega}{\nu}}$  denotes the Womersley number and  $J_o$  is the zeroth order Bessel function of complex argument.

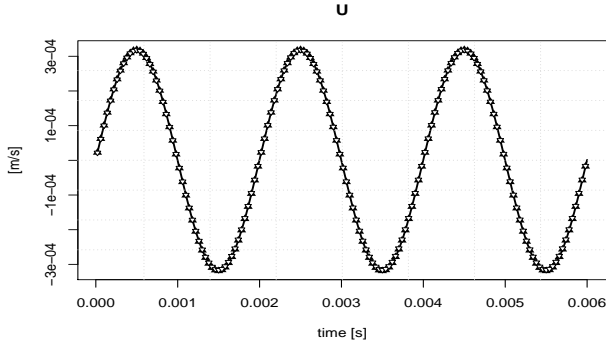
For this test we consider  $L = 1$  m,  $R = 4 \cdot 10^{-3}$  m,  $p_{out} = 5 \cdot 10^5$  Pa,  $A_w = 1000$  Pa,  $f = 500$  Hz and  $\rho_0 = 998.2$  kg/m<sup>3</sup>. According to the solution derived by Womersley, the convective terms are neglected. The numerical parameter for the grids are  $N_x = 500$  and  $N_r = 50$  while the time step has been imposed as  $\Delta t = f^{-1}/100$ , hence a single wave period is resolved in 100 time steps (100 points per wave length). As initial conditions, in the code for the semi-implicit two dimensional method we imposed the velocity profile given by  $u(r, 0)$  while for the one dimensional semi-implicit method, the mean velocity is equal to  $U = 1/A \int u(r, 0) dA$ .

The velocity computed by the scheme SI2D has to be cross-sectionally averaged in order to have a consistent comparison against the velocity given by averaged SI1D method, for which we have implemented and tested several unsteady friction models. Similarly, also the exact solution in Eq. (3.70) is averaged over the cross section, in order to have the exact mean velocity that is subsequently compared against all numerical results.

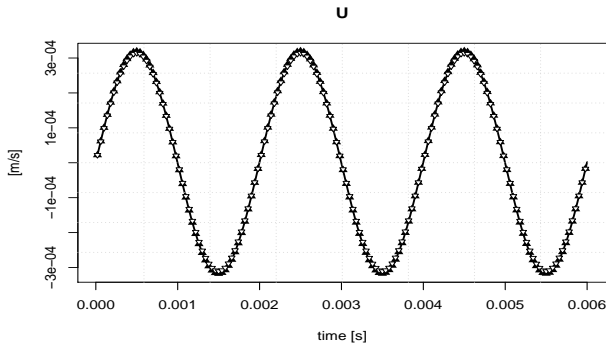
First, a dynamic viscosity of  $\mu = 10^{-3}$  Pa·s is chosen. In Figures 2.4-2.7 the reference solution is well fitted by the numerical solution of SI2D but also by all the friction models implemented in SI1D. Later on, the viscosity has been increased by three order of magnitude, hence  $\mu = 1$  Pa·s. In this case, the exact solution is well fitted by all friction models, apart from the models of Trikha and Brunone, see Fig. 2.9. The first disagreement could be justified by the fact that Trikha's weighting function doesn't approximate well the original function of Zielke. The second mismatch could be explained because the Brunone model does not take into account the past acceleration of the flows, such as the model of Zielke and, consequently, also all its approximations.

### 2.2.2 Impedance matrix for weakly compressible flow in the frequency domain

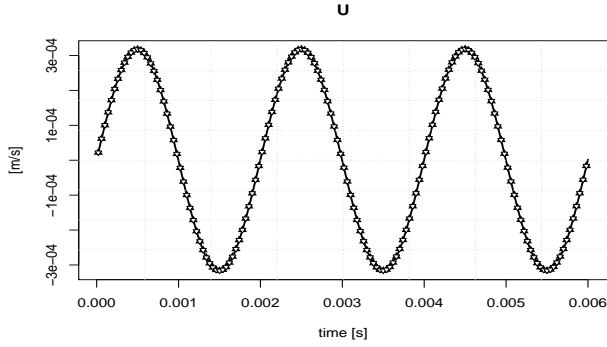
The next investigation regards the comparison of the numerical solutions against an analytic solution derived in the frequency domain that is valid for weakly compressible laminar flows and which has been exposed in [105]. The system of equations (2.27) is simplified neglecting the radial velocity



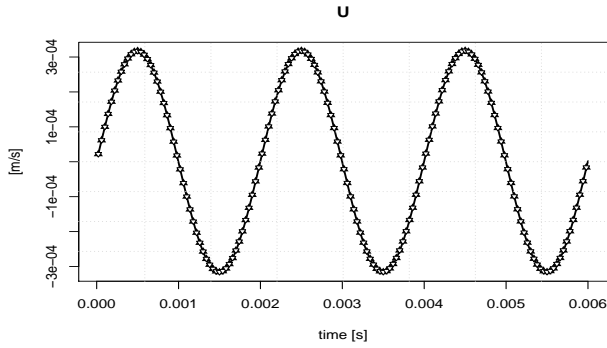
**Figure 2.4:** Womersley test,  $\mu = 10^{-3}$ . Exact solution (—), numerical data given by SI2D (▲) and by SI1D with Zielke model (▼)



**Figure 2.5:** Womersley test,  $\mu = 10^{-3}$ . Exact solution (—), numerical data given by SI1D with Trikha model (▲) and with Brunone model (▼)

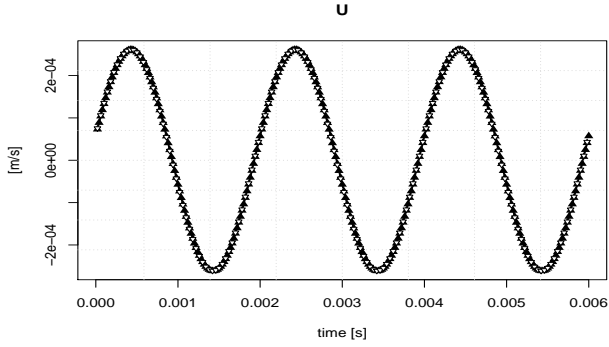


**Figure 2.6:** Womersley test,  $\mu = 10^{-3}$ . Exact solution (—), numerical data given by SIID with Kagawa model ( $\blacktriangle$ ) and with Kagawa ODE model ( $\nabla$ )

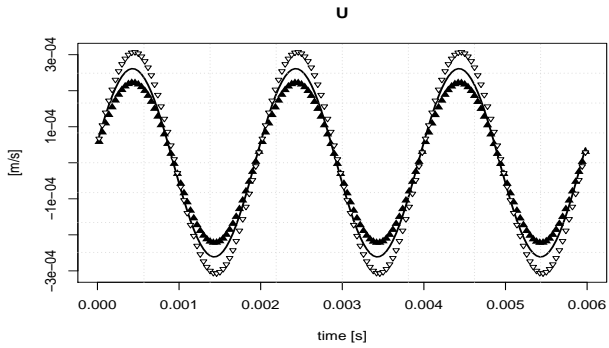


**Figure 2.7:** Womersley test,  $\mu = 10^{-3}$ . Exact solution (—), numerical data given by SIID with UZ model ( $\blacktriangle$ ) and with UZ ODE model ( $\nabla$ )

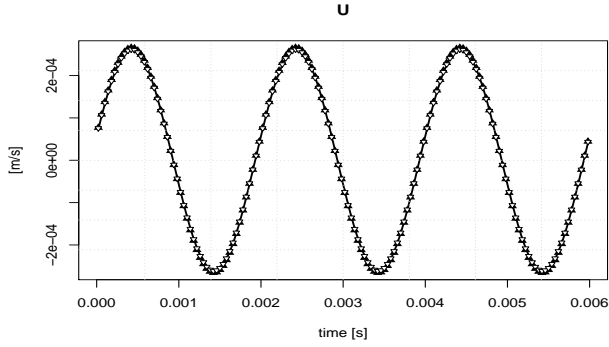




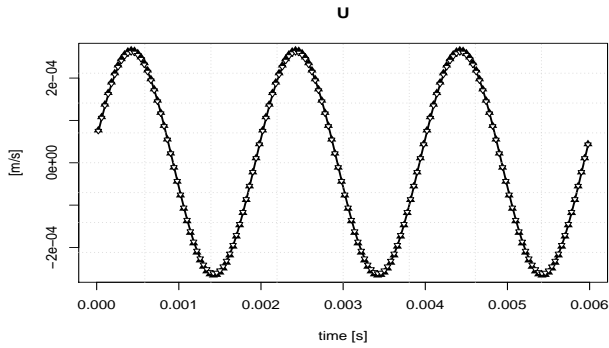
**Figure 2.8:** Womersley test,  $\mu = 1$ . Exact solution (—), numerical data given by SI2D (▲) and by SI1D with Zielke model (▼)



**Figure 2.9:** Womersley test,  $\mu = 1$ . Exact solution (—), numerical data given by SI1D with Trikha model (▲) and with Brunone model (▼)



**Figure 2.10:** Womersley test,  $\mu = 1$ . Exact solution (—), numerical data given by SIID with Kagawa model ( $\blacktriangle$ ) and with Kagawa ODE model ( $\nabla$ )



**Figure 2.11:** Womersley test,  $\mu = 1$ . Exact solution (—), numerical data given by SIID with UZ model ( $\blacktriangle$ ) and with UZ ODE model ( $\nabla$ )

$w$  and the convective terms of the momentum equations. Further assumption is the hypothesis of periodic laminar flow in a situation such that the friction effects given by the frequency are higher than those of quasi-steady flow conditions. Applying the Laplace transform one obtains the following expression for an hydraulic system composed by a tube of length  $L$  and radius  $R$

$$\begin{bmatrix} \hat{p}_1(s) \\ \hat{p}_2(s) \end{bmatrix} = \mathbf{Z}(s) \begin{bmatrix} \hat{Q}_1(s) \\ \hat{Q}_2(s) \end{bmatrix} = \begin{bmatrix} Z_L(s) \frac{\cosh(\gamma(s)L)}{\sinh(\gamma(s)L)} & -Z_L(s) \frac{1}{\sinh(\gamma(s)L)} \\ Z_L(s) \frac{1}{\sinh(\gamma(s)L)} & -Z_L(s) \frac{\cosh(\gamma(s)L)}{\sinh(\gamma(s)L)} \end{bmatrix} \begin{bmatrix} \hat{Q}_1(s) \\ \hat{Q}_2(s) \end{bmatrix}, \quad (2.48)$$

where  $s = j\omega$  is the Laplace parameter,  $\hat{p}(x, s)$  and  $\hat{Q}(x, s)$  are the complex pressure and the complex volume flow rate. The indices 1 and 2 denote the quantities at the inlet and at the outlet of the tube, respectively. Consequently, we have  $\hat{p}_1 = \hat{p}(0, s)$ ,  $\hat{p}_2 = \hat{p}(L, s)$  and  $\hat{Q}_1 = \hat{Q}(0, s)$  and  $\hat{Q}_2 = \hat{Q}(L, s)$ .  $\mathbf{Z}(s)$  is the impedance matrix and its elements are expressed as functions of the hyperbolic sinus and cosine functions. Moreover,  $Z_L$  is a transfer coefficient given by

$$Z_L(s) = Z_0 \sqrt{\frac{J_0(R^*)}{J_1(R^*)}} \quad \text{with} \quad Z_0 = \frac{\sqrt{E'\rho}}{\pi R^2}, \quad E' = E \frac{1}{1 + \frac{E}{E_\infty} W}, \quad (2.49)$$

where  $Z_0$  is the frictionless case of  $Z_L$  and  $E'$  is the modified total bulk modulus of the system that takes into account also the wall elasticity. Here,  $E_\infty$  is the Young modulus of the wall material and  $E$  is the bulk modulus of the fluid, commonly defined as  $E = \rho_0 c_0^2$ ;  $\gamma$  is a wave propagation parameter that is essentially a function of the speed of sound and of the complex radius  $R^*$ :

$$\gamma(s) = \frac{s}{c_0} \sqrt{-\frac{J_0(R^*)}{J_2(R^*)}} \quad \text{with} \quad R^* = j \sqrt{\frac{s}{\nu}} R. \quad (2.50)$$

In order to compute the elements of the impedance matrix for a numerical simulation in the time domain, a frequency has to be fixed and then an oscillatory pressure gradient is imposed, exactly like in the Womersley test. The complex quantities  $\hat{p}$  and  $\hat{Q}$  are then computed by applying the fast Fourier transformation to the time signals of the numerical pressure and mass flow rates taken at the inlet and at the outlet of the tube. Applying an oscillating pressure signal of the same form given by Eq. (3.69) the complex pressure at the outlet  $\hat{p}_2$  is much smaller than the complex pressure at the inlet of the pipe  $\hat{p}_1$ . Moreover, it is possible to demonstrate that the impedance matrix is symmetric and consequently the elements  $z_{11}$  and  $z_{12}$  of the impedance matrix can be computed as follows [79, 105]:

$$z_{11}(s) = \frac{\hat{p}_1(s)\hat{Q}_1(s)}{\hat{Q}_1^2(s) - \hat{Q}_2^2(s)} \quad z_{12}(s) = -\frac{\hat{p}_1(s)\hat{Q}_2(s)}{\hat{Q}_1^2(s) - \hat{Q}_2^2(s)}. \quad (2.51)$$

For additional details about the impedance theory see [105].

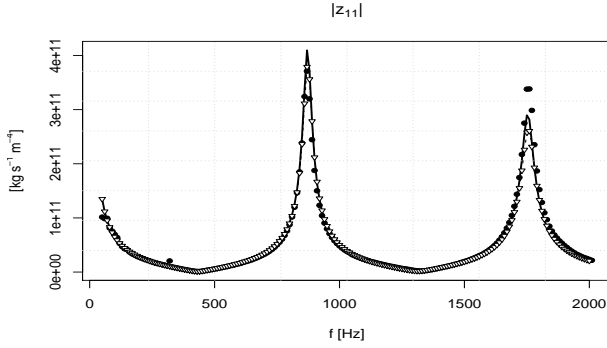
### 2.2.2.1 Rigid case

First, the numerical models are applied in order to reproduce existing experimental measurements that have been obtained for a test fluid in a rigid pipe characterized by radius  $R = 4$  mm and length  $L = 0.791$  m. In the semi-implicit schemes the longitudinal resolution is equal to 1 mm and consequently the number of cells is 791. The radial grid for SI2D is composed by 50 equally spaced layers. On the contrary, for the path conservative method the number of cells  $N_x$  has been chosen equal to 395 that gives a resolution of 2 mm. This choice for the parameters  $\Delta x$  and  $\Delta t$  comes

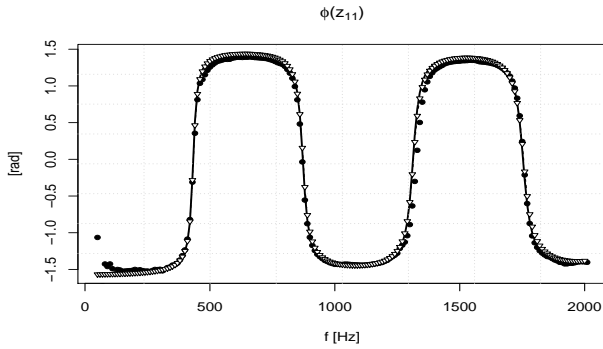
from the result of a series of preliminary simulations that will be discussed in subsection 2.2.4. The investigated frequency range goes from 50 Hz to 2000 Hz using a frequency step of 10 Hz. In addition, both in the numerical simulations as well as in the measured data, a transient phase occurs at the beginning of the experiments. During this period the mass flow signals are not regular, but experiences have shown that they becomes cleaner after 0.1 s. Moreover, the fast Fourier transform has been applied to the signal composed by the last 50 wave periods. Consequently, for each frequency the final time of the simulation is equal to  $0.1 + 50/f$ . According to the hypothesis of this theory the convective terms are neglected and the semi-implicit schemes are unconditionally stable but, in order to limit the numerical viscosity and to resolve the unsteady flow features properly, the parameter  $\Delta t$  is imposed equal to  $T_w/100$ , where  $T_w$  is the wave period ( $T_w = f^{-1}$ ). In Fig. 2.12 one can see that the numerical solution of the scheme S12D reproduces perfectly well the analytic solution without the use of any friction model, since the 2D model is based on first principles and directly discretizes the viscous terms from the available radial velocity profiles. In Fig. 2.13 the first element of the impedance matrix is computed using the one dimensional schemes PC1D and S11D without any unsteady friction model; here, the numerical data do not fit the exact solution and, consequently it is possible to understand the crucial role of unsteady (frequency-dependent) friction models. According to the results of the previous test, also in this case the approximations proposed by Trikha (see Fig. 2.14) and the model of Brunone (see Fig. 2.19) do not give good agreements with the reference solution. In the next four figures (Fig. 2.15, Fig. 2.16, Fig. 2.17 and Fig.2.18) the coefficients of the weighting functions of Kagawa and of Urbanowicz and Kamil are used to compute the contribution of the unsteady wall shear stress by using both the formula of Kagawa (eq. (2.38)) and the new ODE integrator (eq. (3.12)); in all these cases the agreement is very good.

### 2.2.2.2 Elastic case

Second, the frequency domain is investigated for a generic fluid in a flexible pipe. The wall deformation is governed by the Hooke law presented at the beginning of the chapter. It is a very simple model for the tube wall and for comparison against experimental data more sophisticated equations must be used. However, the advantage of the Hook model used for these simulations is that it can be easily taken into account in the exact solution of the impedance matrix. The fluid properties are:  $\rho_0 = 998.2 \text{ Kg/m}^3$ ,  $c_0 = 1400 \text{ m/s}$  and  $\mu = 10^{-2} \text{ Pa}\cdot\text{s}$ . The length of the pipe is  $L = 1 \text{ m}$  and the radius  $R = 1 \text{ mm}$  while the Young's modulus is  $E_\infty = 1 \cdot 10^{10} \text{ Pa}$ . The longitudinal resolution is still 1 mm for the semi-implicit schemes while it is equal to 2 mm for the path conservative method. The other numerical parameters are taken equal to the case of the rigid pipe. The semi-implicit 2D model gives again a very good match compared against the reference solution (see Fig. 2.20) and also for this test when the unsteady friction is neglected in the 1D models the disagreement remains high (see Fig. 2.21). The 1D schemes implemented accounting for the unsteady friction via the convolution integral family match very well the reference solution in the frequency domain (see Figures 2.22, 2.23, 2.24, 2.25 and 2.26). Also the Trikha method works well (see Fig. 2.22) because the viscosity of the fluid is quite small. However, as shown in Fig. 2.27, the model of Brunone still gives significant disagreements when compared against the exact solution in the frequency domain. We want to underline that the approximations of the Zielke model fit the analytic solution for compressible flow in elastic ducts very well, although the original model was proposed for *incompressible* fluids in *rigid* pipes! Moreover, for these tests the simulations are very long and it is practically impossible to get the results by using the original Zielke method due to its

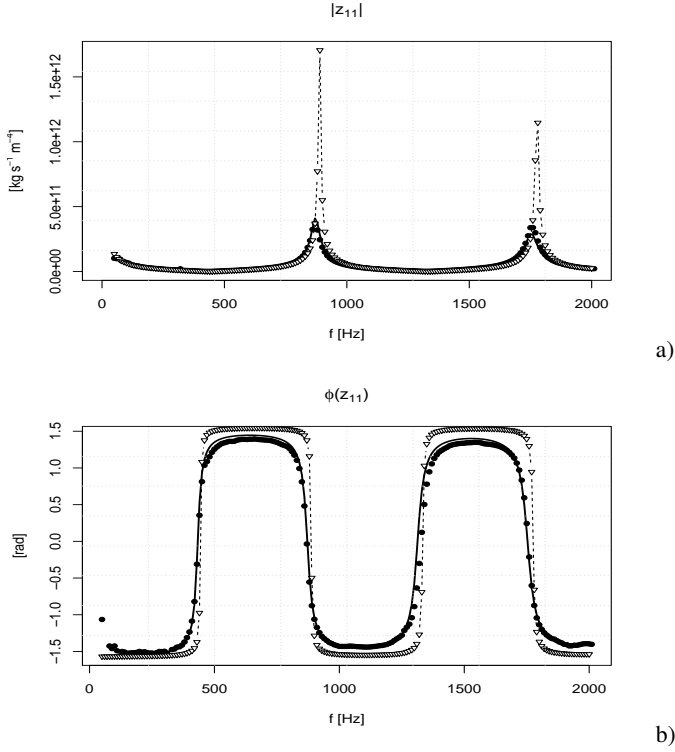


a)

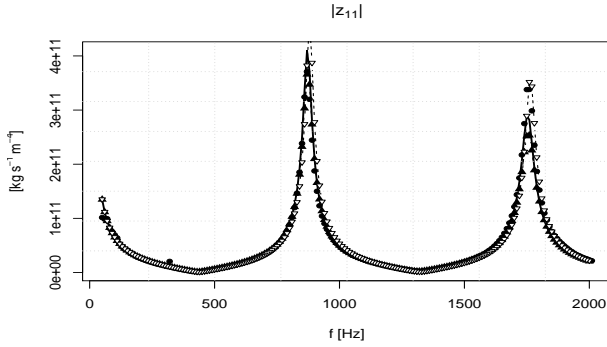


b)

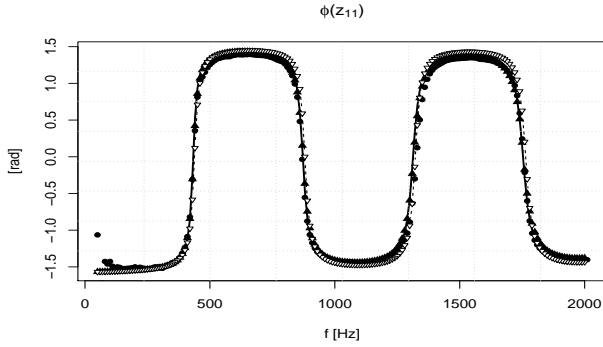
**Figure 2.12:** Absolute value and phase of the element  $z_{11}$  of the impedance matrix for a test fluid in a rigid pipe - Reference solution (—), experimental data (●) and numerical data from the method SI2D ▲



**Figure 2.13:** Absolute value and phase of the element  $z_{11}$  of the impedance matrix for a test fluid in a rigid pipe - Reference solution (—), experimental data (●) and numerical data from the method SI1D (▲) and from PC1D (▽) using only a steady friction model

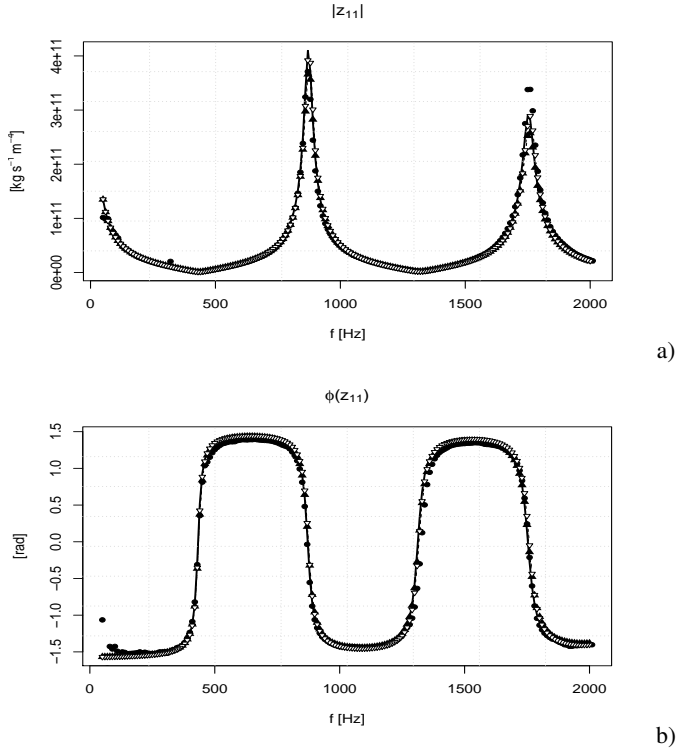


a)



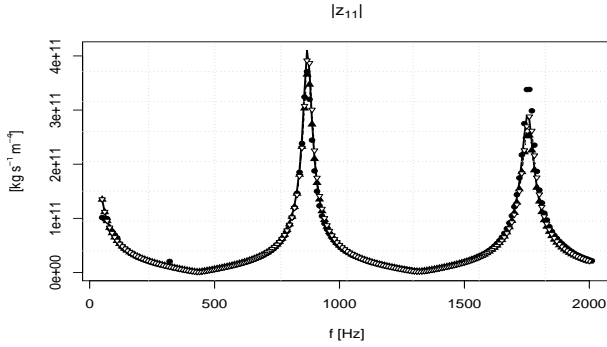
b)

**Figure 2.14:** Absolute value and phase of the element  $z_{11}$  of the impedance matrix for a test fluid in a rigid pipe - Reference solution (—), experimental data (●) and numerical data from the method SIID (▲) and from PC1D (▽) with Trikha model for the unsteady friction term

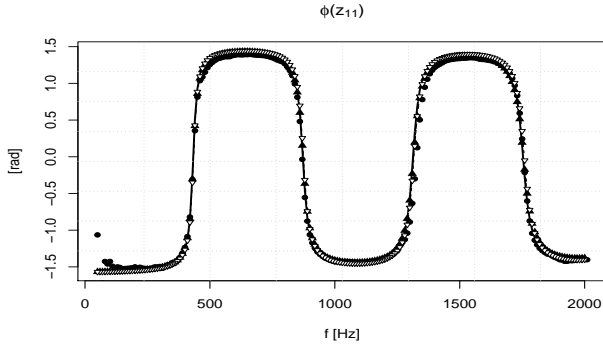


**Figure 2.15:** Absolute value and phase of the element  $z_{11}$  of the impedance matrix for a test fluid in a rigid pipe - Reference solution (—), experimental data (●) and numerical data from the method SI1D (▲) and from PC1D (▽) with Kagawa model for the unsteady friction term



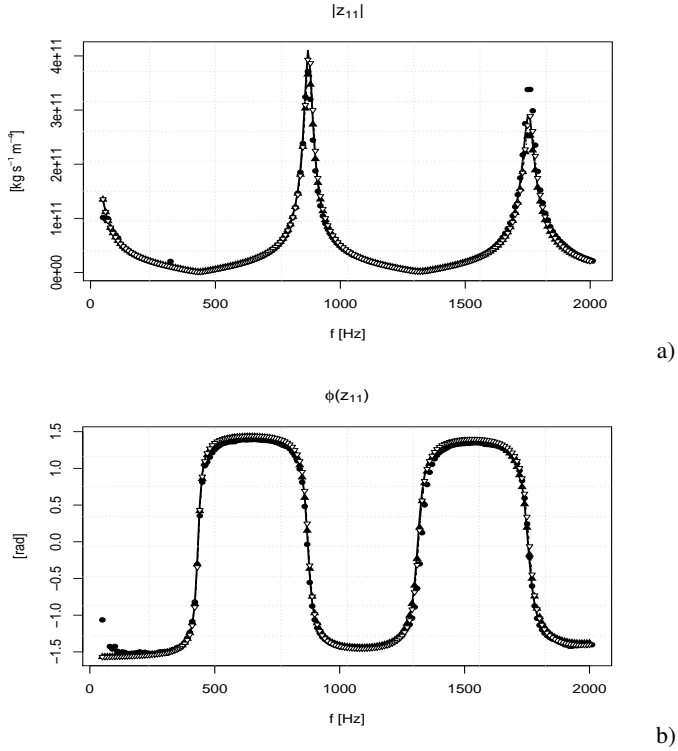


a)

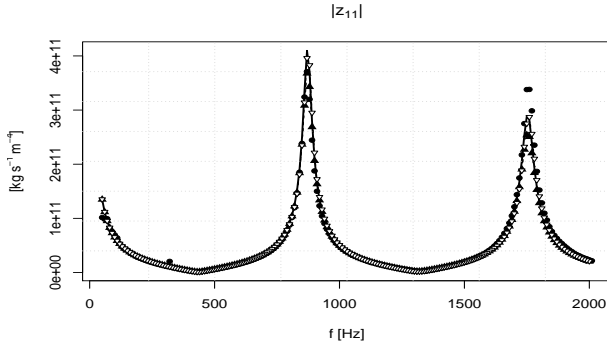


b)

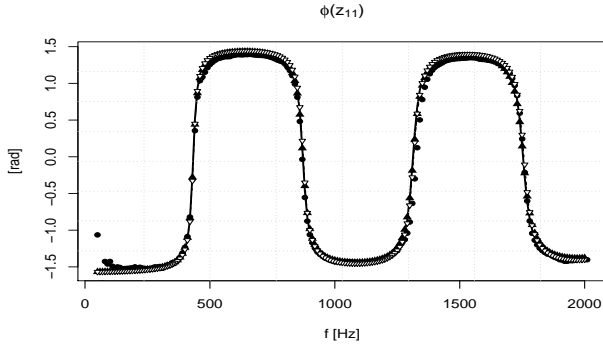
**Figure 2.16:** Absolute value and phase of the element  $z_{11}$  of the impedance matrix for a test fluid in a rigid pipe - Reference solution (—), experimental data (●) and numerical data from the method SIID (▲) and from PC1D (▽) with Kagawa ODE model for the unsteady friction term



**Figure 2.17:** Absolute value and phase of the element  $z_{11}$  of the impedance matrix for a test fluid in a rigid pipe - Reference solution (—), experimental data (●) and numerical data from the method SI1D (▲) and from PC1D (▽) with Urbanowicz and Zarzycky model for the unsteady friction term

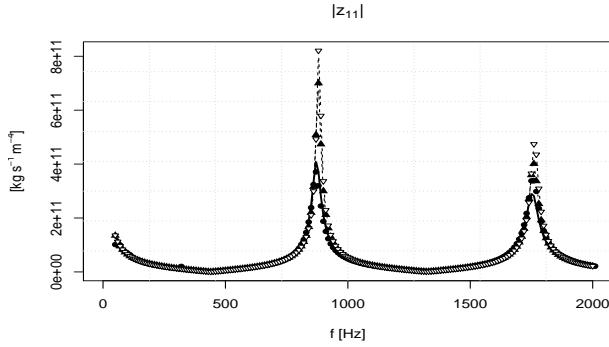


a)

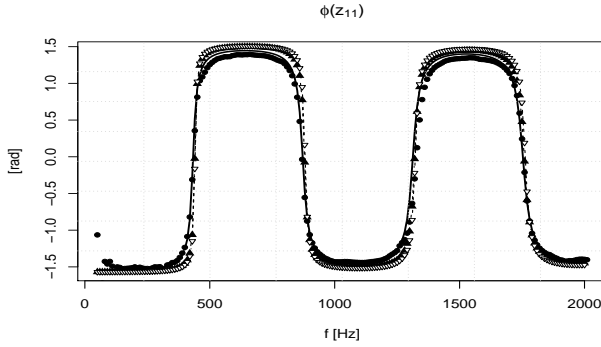


b)

**Figure 2.18:** Absolute value and phase of the element  $z_{11}$  of the impedance matrix for a test fluid in a rigid pipe - Reference solution (—), experimental data (●) and numerical data from the method SIID (▲) and from PC1D (▽) with Urbanowicz and Zarzycky ODE model for the unsteady friction term



a)



b)

**Figure 2.19:** Absolute value and phase of the element  $z_{11}$  of the impedance matrix for a test fluid in a rigid pipe - Reference solution (—), experimental data (●) and numerical data from the method SI1D (▲) and from PC1D (▽) with Brunone model for the unsteady friction term

slowness and due to the massive memory amount that it requires.

### 2.2.3 Water hammer

The last test presented here is a typical benchmark problem used to assess the quality of different unsteady friction models. In a long pipeline characterized by normal flow and with a constant hydraulic head upstream given for example by a reservoir, the downstream valve is suddenly closed generating a pressure wave that travels back into the pipe and then it is periodically reflected. This is known as the experiment of Holmboe and Rouleau [89] and from their results it was understood that the unsteady friction contribution is crucial in order to predict the pressure damping at the downstream border of the pipe correctly. Here, we consider a fluid which has similar physical properties of water:  $\rho_0 = 998.2 \text{ kg/m}^3$ ,  $c_0 = 1400 \text{ m/s}$ ,  $\mu = 10^{-3} \text{ Pa}\cdot\text{s}$ . The length of the pipe is  $L = 15 \text{ m}$  and the radius is  $R = 4 \text{ mm}$ . For all the schemes the number of cells is  $N_x = 1000$ , because we decided to take the pressure signals in the same position in order to compare the numerical data in the best way possible. In addition, for SI2D the number of rings is  $N_r = 50$  and for both the semi-implicit schemes the time step  $\Delta t$  is imposed equal to  $2 \cdot 10^{-5} \text{ s}$  to reduce the numerical viscosity introduced by the implicit time discretization.

The aim of this analysis is to compare the behavior of the different unsteady friction models considered in this chapter. We assume that the reference solution for the pressure at the end of the tube is the numerical one given by SI2D, since it is the only method that is able to evaluate the friction term directly from first principles, without any closure relation.

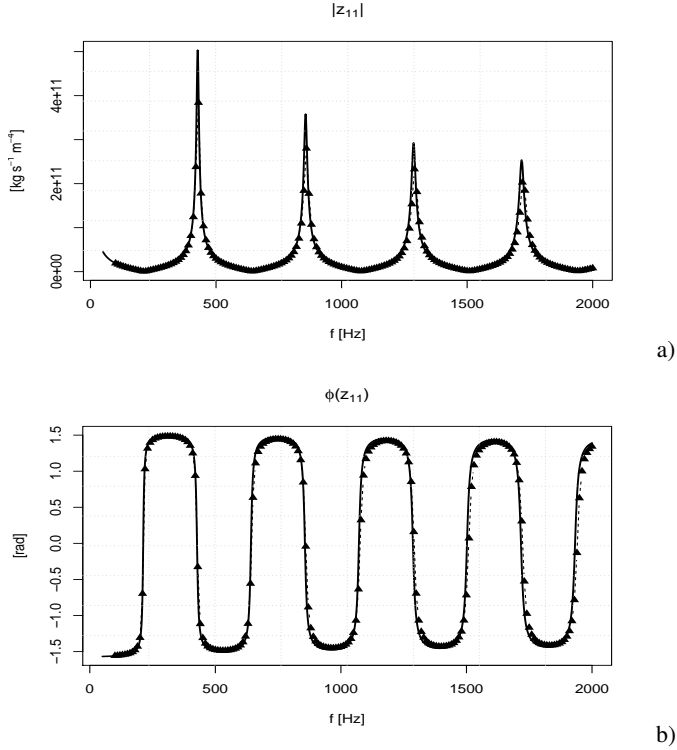
In Fig. 2.28 the results of the Trikha method implemented in SI1D and in PC1D are shown. It is possible to observe that both numerical schemes (SI1D and PC1D) give very similar results with the same unsteady friction model, which, however, deviate from the reference solution. In Fig. 2.29 the numerical data obtained from the Brunone model are exposed and a disagreement can be seen, but this behavior is intrinsic to the Brunone model and has been observed also by other authors, see for example [120].

Fig. 2.30 and Fig. 2.31 show the results of the Kagawa coefficients used in the original formula of Kagawa (2.38) and in our new ODE method (3.12). The results illustrate a very good fitting with the reference solution. The same conclusion can be observed applying the UZ model coefficients, see Figures 2.32 and 2.33.

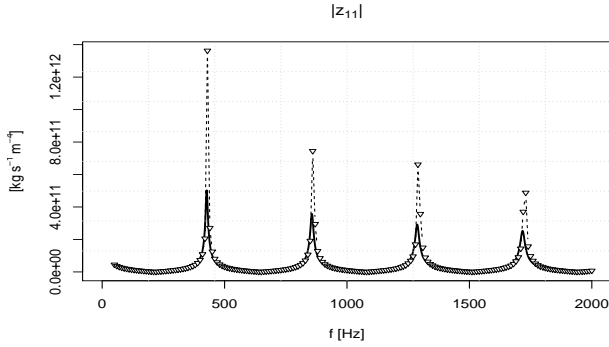
As conclusive test, we repeat the previous simulations for an elastic tube with  $E_\infty = 10^{10} \text{ Pa}$  and with an increased viscosity equal to  $\mu = 10^{-1} \text{ Pa}\cdot\text{s}$ . This comparison in the time domain is carried out in order to show again the robustness of our new ODE method at relatively high viscosities and to underline again the excellent fitting of convolution-integral type unsteady friction models for viscous compressible fluids in flexible pipes. Figures 2.34 and 2.35 show the good agreement between the reference solution and the 1D models implemented with the Kagawa ODE and with the UZ ODE models, respectively.

### 2.2.4 Efficiency analysis and computational times

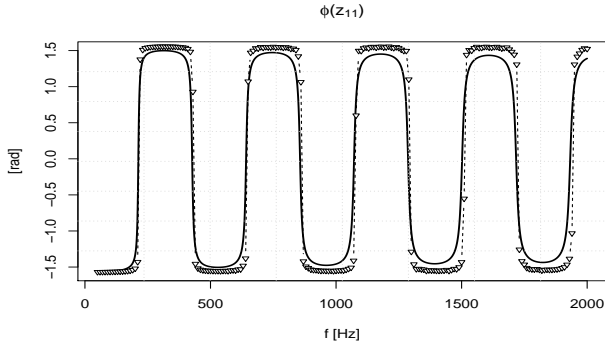
The reference solution for the impedance matrix allows to compute the error and to compare it against the CPU times for the different schemes here considered. The error calculated is intended as the mean of the relative errors between the numerical solution and the analytical one related to



**Figure 2.20:** Absolute value and phase of the element  $z_{11}$  of the impedance matrix for a compressible fluid in an elastic pipe - Reference solution (—) and numerical data from the method SI2D ▲

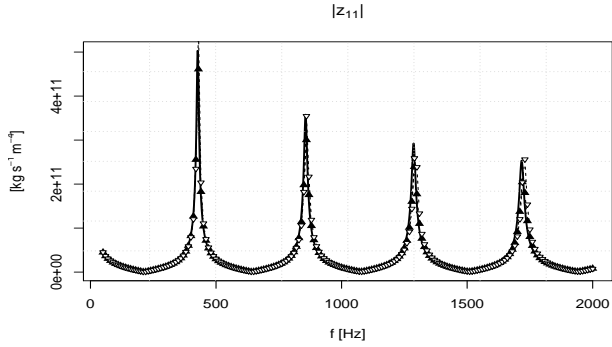


a)

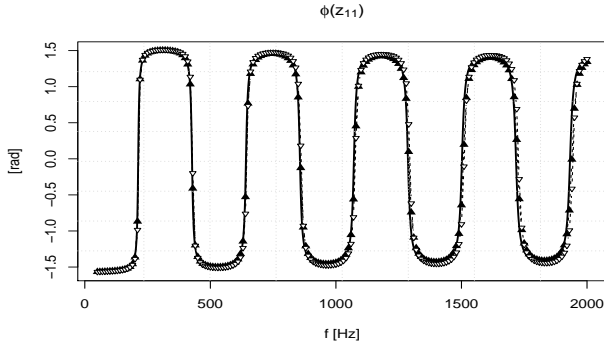


b)

**Figure 2.21:** Absolute value and phase of the element  $z_{11}$  of the impedance matrix for a compressible fluid in an elastic pipe - Reference solution (—) and numerical data from the method SI1D (▲) and from PC1D (▽) accounting only the steady friction term



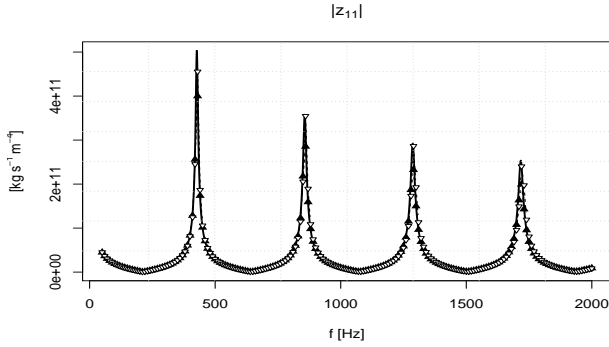
a)



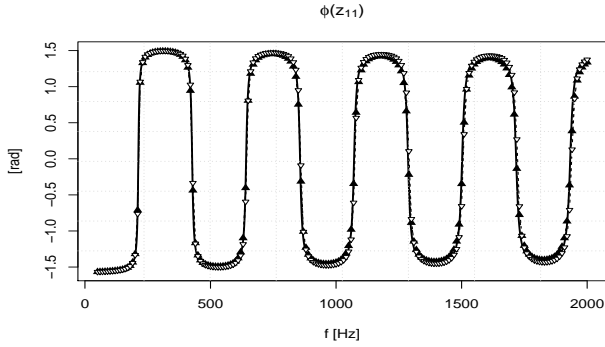
b)

**Figure 2.22:** Absolute value and phase of the element  $z_{11}$  of the impedance matrix for a compressible fluid in an elastic pipe - Reference solution (—) and numerical data from the method SI1D (▲) and from PC1D (▽) with Trikha model for the unsteady friction term



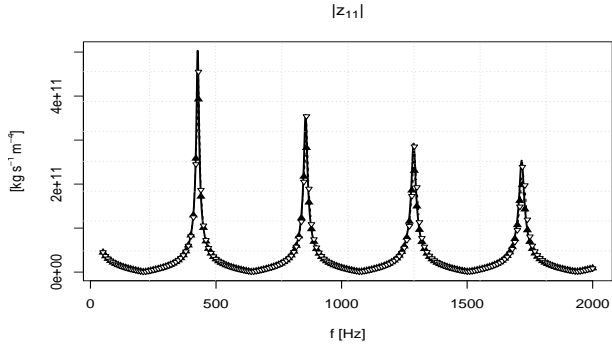


a)

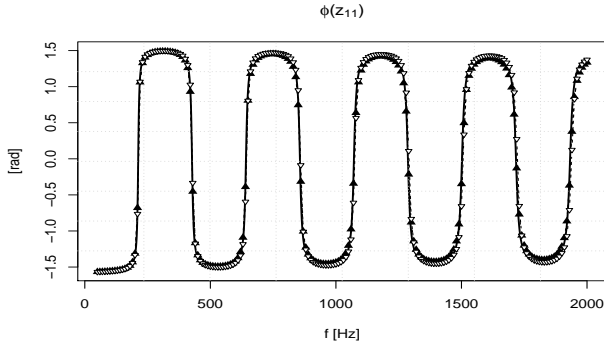


b)

**Figure 2.23:** Absolute value and phase of the element  $z_{11}$  of the impedance matrix for a compressible fluid in an elastic pipe - Reference solution (—) and numerical data from the method SIID ( $\blacktriangle$ ) and from PC1D ( $\nabla$ ) with Kagawa model for the unsteady friction term

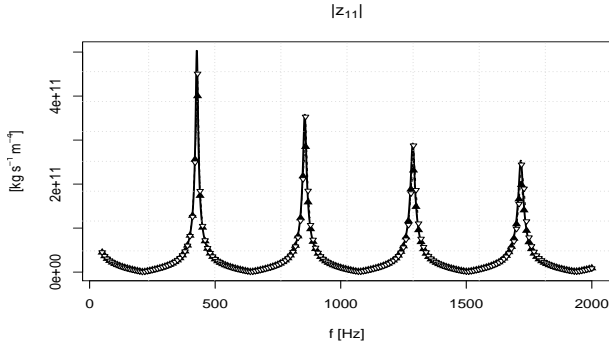


a)

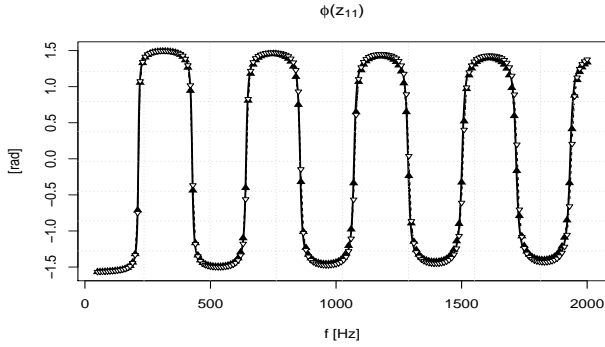


b)

**Figure 2.24:** Absolute value and phase of the element  $z_{11}$  of the impedance matrix for a compressible fluid in an elastic pipe - Reference solution (—) and numerical data from the method SI1D ( $\blacktriangle$ ) and from PC1D ( $\nabla$ ) with Kagawa ODE model for the unsteady friction term

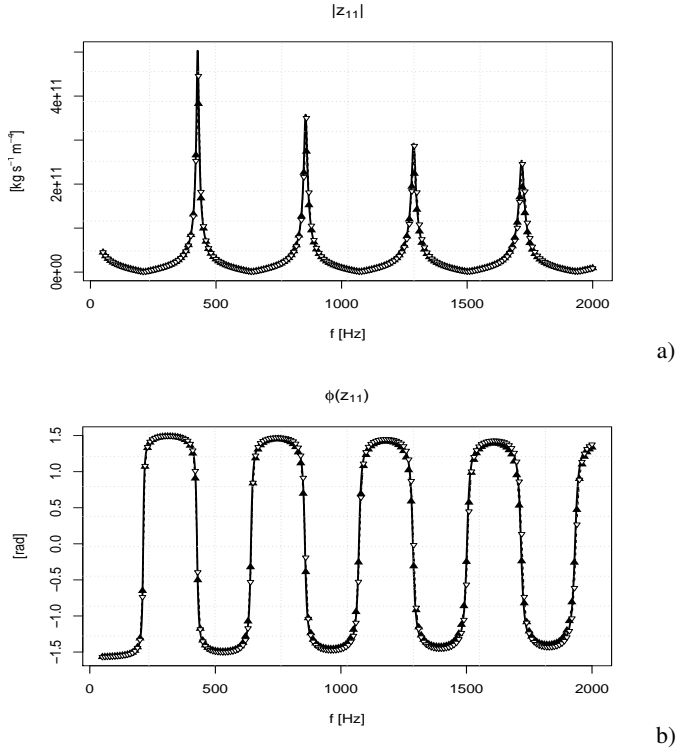


a)

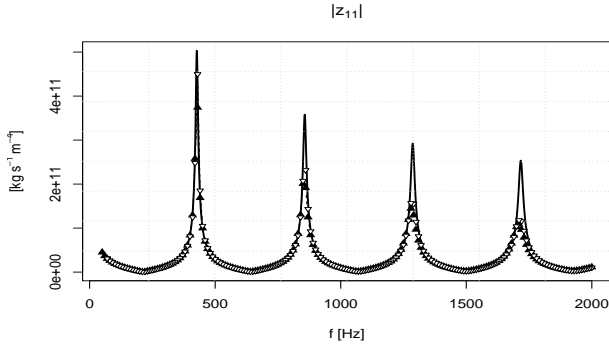


b)

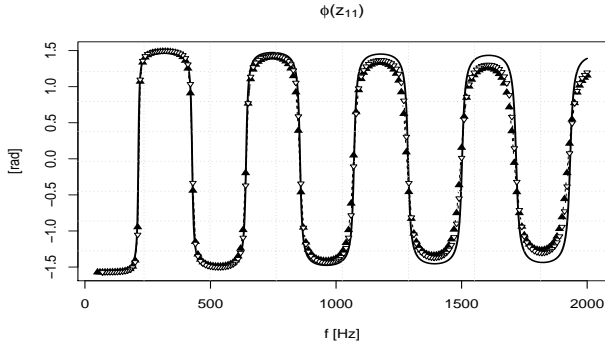
**Figure 2.25:** Absolute value and phase of the element  $z_{11}$  of the impedance matrix for a compressible fluid in an elastic pipe - Reference solution (—) and numerical data from the method SIID ( $\blacktriangle$ ) and from PC1D ( $\nabla$ ) with Urbanowicz and Zarzycky model for the unsteady friction term



**Figure 2.26:** Absolute value and phase of the element  $z_{11}$  of the impedance matrix for a compressible fluid in an elastic pipe - Reference solution (—) and numerical data from the method SI1D ( $\blacktriangle$ ) and from PC1D ( $\nabla$ ) with Urbanowicz and Zarzycky ODE model for the unsteady friction term

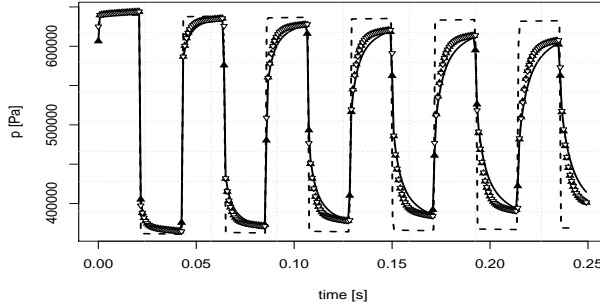


a)

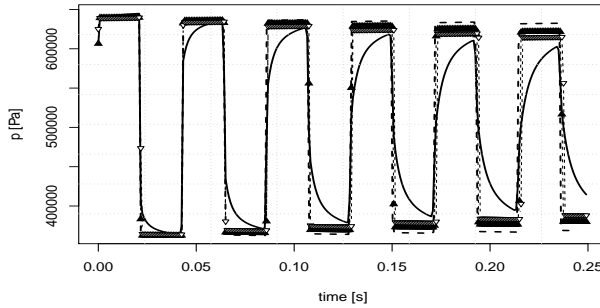


b)

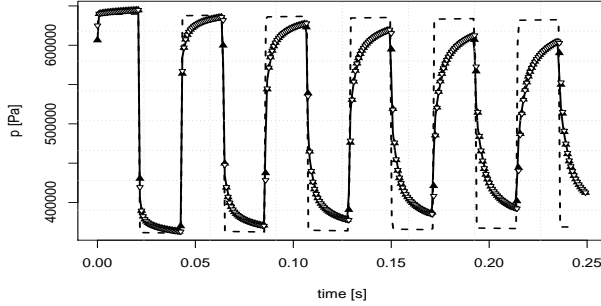
**Figure 2.27:** Absolute value and phase of the element  $z_{11}$  of the impedance matrix for a compressible fluid in an elastic pipe - Reference solution (—) and numerical data from the method SI1D ( $\blacktriangle$ ) and from PC1D ( $\nabla$ ) with Brunone model for the unsteady friction term



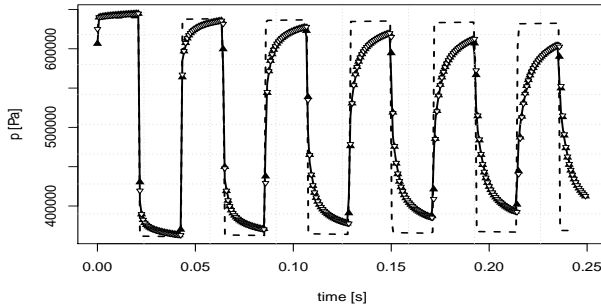
**Figure 2.28:** Water hammer, rigid pipe - Reference solution (—), 1D methods without unsteady friction (- - -), numerical data given by SI1D with Trikha model (▲) and by PC1D with Trikha model (▽)



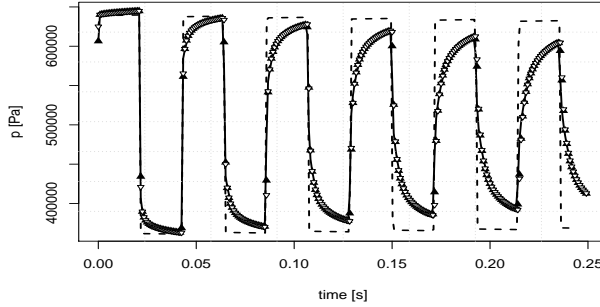
**Figure 2.29:** Water hammer, rigid pipe - Reference solution (—), 1D methods without unsteady friction (- - -), numerical data given by SI1D with Brunone model (▲) and by PC1D with Brunone model (▽)



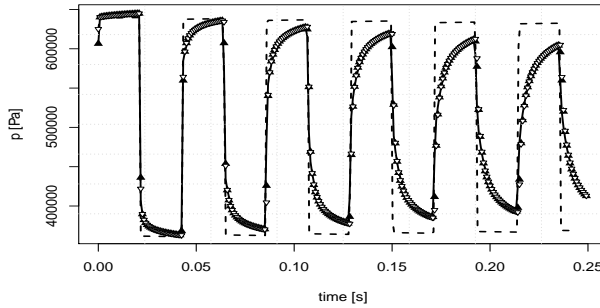
**Figure 2.30:** Water hammer, rigid pipe - Reference solution (—), 1D methods without unsteady friction (---), numerical data given by SI1D with Kagawa model (▲) and by PC1D with Kagawa model (▽)



**Figure 2.31:** Water hammer, rigid pipe - Reference solution (—), 1D methods without unsteady friction (---), numerical data given by SI1D with Kagawa ODE model (▲) and by PC1D with Kagawa ODE model (▽)

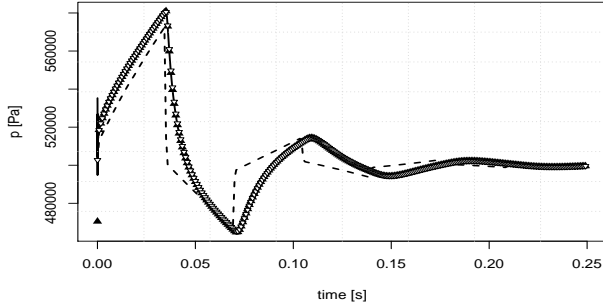


**Figure 2.32:** Water hammer, rigid pipe - Reference solution (—), 1D methods without unsteady friction (---), numerical data given by SI1D with UZ model (▲) and by PC1D with UZ model (▽)

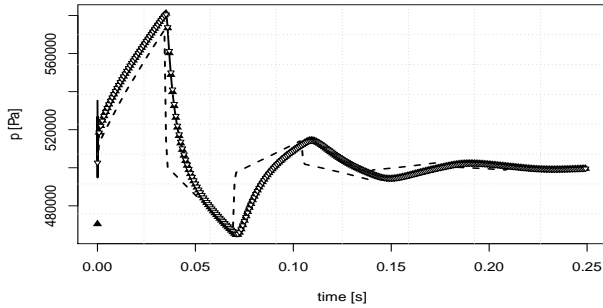


**Figure 2.33:** Water hammer, rigid pipe - Reference solution (—), 1D methods without unsteady friction (---), numerical data given by SI1D with UZ ODE model (▲) and by PC1D with UZ ODE model (▽)





**Figure 2.34:** Water hammer, elastic pipe - Reference solution (—), 1D methods without unsteady friction (---), numerical data given by SI1D with Kagawa ODE model (▲) and by PC1D with Kagawa ODE model (▽)



**Figure 2.35:** Water hammer, elastic pipe - Reference solution (—), 1D methods without unsteady friction (---), numerical data given by SI1D with UZ ODE model (▲) and by PC1D with UZ ODE model (▽)

## 2 Explicit and semi-implicit FV schemes for compressible flows in elastic pipes

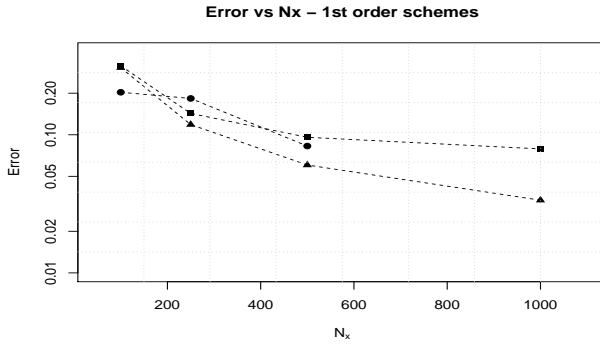
the absolute value of the complex parameter  $z_{11}$ :

$$\epsilon_{|z_{11}|} = \frac{1}{N_f} \sum_{i=1}^{N_f} \left| \frac{|z_{11}^{num}(f_i)| - |z_{11}^{ex}(f_i)|}{|z_{11}^{ex}(f_i)|} \right| . \quad (2.52)$$

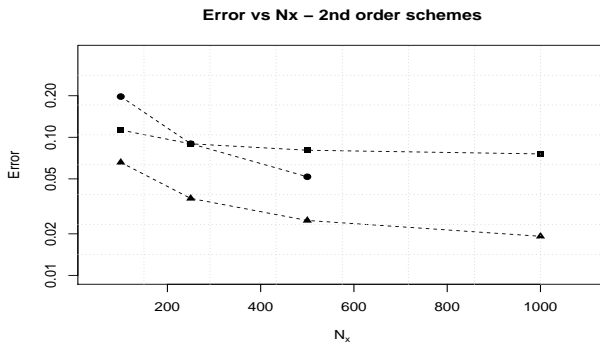
This analysis has been carried out for a test fluid characterized by  $c_0 = 1400$  m/s,  $\rho_0 = 998.2$  kg/m<sup>3</sup> and  $\mu = 10^{-2}$  Pa/s in a rigid pipe of length  $L = 1$  m. Here, the friction model used in the 1D models is the ODE method with the coefficients of Kagawa. First, we consider the error of all first order schemes with three different meshes with  $N_x = 100, 250, 500$ ; for the explicit schemes the time step size  $\Delta t$  is given by the CFL condition with CFL=0.9, while for semi-implicit schemes it is chosen as  $\Delta t = T/N_x$ , so that a wave period is resolved with  $N_x$  time steps (points per wave length). The results of this analysis are given in Table 2.1. In Fig. 2.36 it is possible to see that when the number of longitudinal cells increases, the errors decrease. We then have repeated all the simulations using the second order schemes and from Tab. 2.1 and from the plot in Fig 2.37 it can be observed that the error is much lower than in the first order case, especially for the semi-implicit schemes. Furthermore, in some cases for SI2D the error is also lower than the one of the second order explicit schemes. Moreover, for the semi-implicit schemes we decided to introduce also a grid where  $N_x$  is equal to 1000, because experiments have shown that the qualitative solution improved in particular where the parameters  $|z_{11}|$  and  $|z_{12}|$  have peaks. At the end of this analysis, we decided to simulate the frequency domain with a resolution of 2 mm for the explicit methods, while for the semi-implicit schemes we have chosen  $\Delta x = 1$  mm and with a time resolution of  $\Delta t = T/100$ .

Furthermore, the CPU times of the simulations have been plotted against the errors both for the 1st order and for the 2nd order schemes (see Fig.2.38 and Fig. 2.39). In general, for a given error the semi-implicit 1D method gives results in much less time with respect to the explicit 1D scheme and w.r.t. SI2D. For the second order case, SI2D is computationally more efficient than the explicit schemes for a given spatial grid. The computational times of the simulations done for the impedance matrix and for the water hammer have been measured and they are exposed in Tab. 2.3 for SI2D and for the one dimensional schemes.

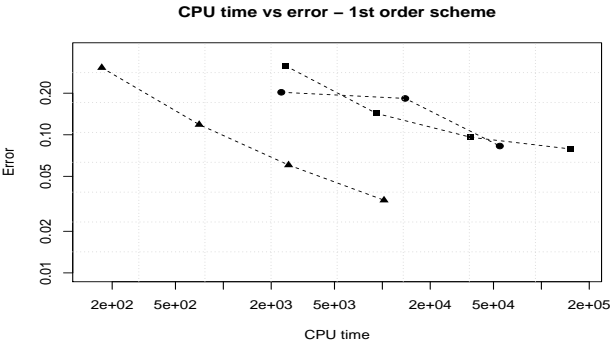
It is possible to observe that for the impedance matrix for the rigid tube the CPU time of the scheme SI2D is very similar to the one of the explicit 1D methods both with the Osher RS and with the HLL RS. However, for the elastic case the 1D explicit method becomes extremely slow (due to the necessary iterative computation of the primitive variables from the vector of conserved quantities), and one can observe that the two dimensional *semi-implicit* method is almost one order of magnitude *faster* than the *explicit* 1D methods, which is quite a remarkable result. Moreover, the experiments done using the Osher RS have produced the results in less time with respect to the HLL RS. This last fact occurs because the iterative procedure for the calculation of the pressure ( $g(p) = Q_1 - \rho(p)A(p)$ ) is required more times in the computation of the HLL flux than in the evaluation of the Osher flux. We think this is a very particular case for this non-linear system of PDEs studied in this work. For systems composed by more than 2 PDEs the Osher RS is in general more expensive. Anyway, for all the test cases the method SI1D is clearly the fastest of the schemes that have been considered here, since it represents a special case of SI2D when  $N_r = 1$ ; it gives very accurate results in a very small computational time. Last but not least, in all simulations we observe that our new ODE integrator method (3.12) for the Zielke convolution integral is faster than the original formula of Kagawa. This is valid using both, the Kagawa coefficients as well as the Urbanowicz and Zarzycky coefficients. We attribute this result to the fact that the new ODE method does not involve time-consuming evaluations of exponential functions.



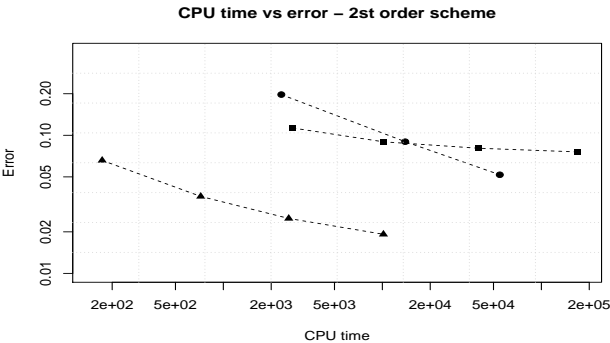
**Figure 2.36:**  $N_x$  and errors for the first order schemes - ● Explicit schemes, ▲ SI1D, ■ SI2D



**Figure 2.37:**  $N_x$  and errors for the second order schemes - ● Explicit schemes, ▲ SI1D, ■ SI2D



**Figure 2.38:** CPU times and errors for the first order schemes - ● Explicit schemes, ▲ SI1D, ■ SI2D



**Figure 2.39:** CPU times and errors for the second order schemes - ● Explicit schemes, ▲ SI1D, ■ SI2D

Explicit Scheme				
$N_x$	$\Delta x$ [mm]	$\epsilon_{ z_{11} }$	1st order	$\epsilon_{ z_{11} }$ 2nd order
100	10		0.2031	0.1972
250	4		0.1833	0.0899
500	2		0.0529	0.0518
1000	1		-	-
Semi-implicit 1D				
$N_x$	$\Delta x$ [mm]	$\epsilon_{ z_{11} }$	1st order	$\epsilon_{ z_{11} }$ 2nd order
100	10		0.3061	0.0658
250	4		0.1185	0.0361
500	2		0.0604	0.0250
1000	1		0.0336	0.0192
Semi-implicit 2D				
$N_x$	$\Delta x$ [mm]	$\epsilon_{ z_{11} }$	1st order	$\epsilon_{ z_{11} }$ 2nd order
100	10		0.3157	0.1129
250	4		0.1428	0.0897
500	2		0.0959	0.0806
1000	1		0.0790	0.0758

**Table 2.1:** Frequency domain - error  $\epsilon_{|z_{11}|}$  for the first order schemes and for the second order schemes corresponding to different  $N_x$

## 2.3 Conclusions concerning the comparison of explicit versus semi-implicit schemes for compressible pipe flow

In this chapter we have discussed, implemented and compared several finite volume schemes for systems of partial differential equations that govern pressurized viscous compressible flows in rigid and flexible pipes. Since the governing PDE system contains a non-conservative product, the classical explicit Godunov-type schemes have to be supplemented by using a path conservative approach in order to consider the varying cross section due to the elastic wall deformation. Another family of methods under consideration is the class of staggered semi-implicit finite volume schemes. For this class of methods, the two-dimensional case gives accurate results without the need of any closure for the viscous terms. Furthermore, in most cases the 2D semi-implicit method is almost as fast as the explicit one-dimensional scheme, but in some cases it can even be faster.

Moreover, in order to match the exact solution and the experimental measurements, all one dimensional models need a closure relation for the wall friction, in particular for the unsteady contribution to the wall shear stress that is crucial for the simulations that have been considered here. Several

CPU time [s]		Explicit schemes	
$N_x$	$\Delta x$ [mm]	1st order	2nd order
100	10	1989	2315
250	4	11742	13923
500	2	39810	54776
1000	1	-	-

---

CPU time [s]		Semi-implicit 1D	
$N_x$	$\Delta x$ [mm]	1st order	2nd order
100	10	172	173
250	4	705	721
500	2	2574	2578
1000	1	10246	10158

---

CPU time [s]		Semi-implicit 2D	
$N_x$	$\Delta x$ [mm]	1st order	2nd order
100	10	2462	2741
250	4	9177	10171
500	2	35962	40208
1000	1	151680	169475

**Table 2.2:** Frequency domain - CPU times for the first order schemes and for the second order schemes corresponding to different  $N_x$

unsteady friction models have been incorporated into the 1D schemes. The thorough comparisons have shown that the models of Brunone and of Trikha do not give a good fitting over the entire range of Reynolds numbers studied here, especially when the viscosity assumes large values. The most accurate models seem to be those that approximate the convolution integral of Zielke, such as the model of Kagawa and the one Urbanowicz and Zarzycki. Furthermore, we have proposed a new formula for the evaluation of the convolution integral of Zielke that is at least as accurate as the formula of Kagawa, but it is faster since it does not involve the time consuming evaluation of exponential functions.

### 2.3 Conclusions about explicit vs. semi-implicit schemes for pipe flow

Impedance matrix (Rigid pipe)			
SI2D	50550		
1D	SI1D	PC Osher	PC HLL
Trikha	784	37575	39623
Kagawa	1204	40408	41594
Kagawa ODE	832	38030	39630
UZ	3143	45037	46536
UZ ODE	2503	39988	40145
Brunone	863	37233	40272
Impedance matrix (Elastic pipe)			
SI2D	66059		
1D	SI1D	PC Osher	PC HLL
Trikha	1391	419309	549111
Kagawa	1968	405697	552679
Kagawa ODE	1483	405517	552166
UZ	3387	426222	578988
UZ ODE	2115	422963	583574
Brunone	1352	415219	573813
Water Hammer (Rigid)			
SI2D	227		
1D	SI1D	PC Osher	PC HLL
Trikha	16	121	193
Kagawa	28	135	207
Kagawa ODE	22	127	200
UZ	56	164	239
UZ ODE	38	141	219
Brunone	15	120	193

**Table 2.3:** CPU times in seconds for the two-dimensional method compared with different one dimensional methods





### 3 Semi-implicit staggered discontinuous Galerkin schemes for axially symmetric viscous compressible flows in elastic tubes

All numerical methods presented in the previous chapter were at most second order accurate in space. In this chapter we therefore propose a novel family of staggered semi-implicit discontinuous Galerkin (DG) finite element schemes for the simulation of axially symmetric, weakly compressible and laminar viscous flows in elastic pipes that are, at least in principle, of arbitrary order of accuracy in space. The equation of state (EOS) of the fluid is assumed to be barotropic and two different mathematical models derived from the compressible Navier-Stokes equations are considered in this chapter.

The first model describes cross-sectionally averaged 1D flows, including steady and frequency-dependent wall friction effects. The novelty of our numerical method compared to standard DG schemes consists in the use of a *staggered mesh*, where the pressure is defined over a primary grid and the velocity field is defined on edge-based staggered dual control volumes. This approach is well known from classical semi-implicit finite difference schemes for the incompressible Navier-Stokes equations, but it is still quite unusual for high order DG schemes. The continuity equation is integrated over the control volumes that belong to the main grid, while the momentum equation is integrated over the elements of the edge-based staggered dual grid. The nonlinear convective terms are discretized explicitly, while the pressure gradient and the mass flux are discretized implicitly. Up to second order of accuracy in time can be achieved with the so-called  $\theta$ -method. Inserting the discrete momentum equation in the discrete mass conservation equation leads to a mildly nonlinear algebraic system for the degrees of freedom of the pressure. Such mildly nonlinear systems can be very efficiently solved using the Newton algorithm of Casulli *et al.*. We observe that the linear part of the mildly nonlinear system is symmetric and positive definite.

The second model is derived from the compressible Navier-Stokes equations in cylindrical coordinates. Assuming hydrostatic flow with constant pressure inside each cross section as well as axial symmetry, only the terms in the axial and the radial direction need to be considered. Therefore, we call the second model the  $2D_{xT}$  model. Also in this case we use a staggered mesh for pressure and velocity and thus the same philosophy as for the 1D model can be applied to obtain the discrete pressure system. For the  $2D_{xT}$  model a staggered DG scheme is also applied for the computation of the viscous stress tensor in the discrete momentum equation. However, in radial direction the resulting linear system for the friction terms is not symmetric and is thus solved using the Thomas algorithm for block three-diagonal systems.

The use of a semi-implicit DG scheme leads to a very mild CFL condition based only on the fluid velocity and not on the sound speed, which makes the method very efficient, in particular in the limit cases when the speed of sound of the fluid tends to infinity (incompressible fluid) and in the rigid case where the wall strain of the pipe tends to zero. In addition, at every Newton step a symmetric positive definite and well conditioned block three-diagonal linear system is solved for the pressure, using either a matrix-free conjugate gradient method, or the Thomas algorithm properly extended

to block-tridiagonal systems. Moreover, when the polynomial degree of the basis and test functions is equal to zero the schemes reduce to classical semi-implicit finite volume methods, such as the ones presented in the previous chapter.

While in the  $2D_{xr}$  model the viscous effects in radial direction are directly obtained from first principles via the Navier-Stokes equations, the 1D model requires an additional closure relation for the wall friction. For both models we perform several tests in order to validate the numerical methods for steady and unsteady flows of compressible and nearly incompressible fluids in elastic and rigid tubes. We also provide numerical convergence results in order to show that the developed schemes achieve high order of accuracy in space.

### 3.1 Staggered semi-implicit DG scheme applied to the 1D model

#### 3.1.1 Governing equations of the 1D model

The dynamics of fluids in general time-dependent domains is governed by the compressible Navier-Stokes equations. Supposing axially symmetric flow and that the longitudinal scale is much larger than the radial one, it is possible to simplify the system and to introduce the hypothesis of hydrostatic equilibrium, i.e. the pressure is constant in each cross section. Moreover, we assume laminar flow. Under these assumptions one obtains the following one-dimensional, cross-sectionally averaged model that consists of two PDE, namely the continuity equation and the momentum equation, see [56, 91, 104, 141],

$$\frac{\partial}{\partial t}(\rho A) + \frac{\partial}{\partial x}(\rho AU) = 0, \quad (3.1a)$$

$$\frac{\partial}{\partial t}(\rho AU) + \frac{\partial}{\partial x}(\rho AU^2) + A \frac{\partial p}{\partial x} = -2\pi R\tau_s - 2\pi R\tau_u, \quad (3.1b)$$

where  $t \in \mathbb{R}_0^+$  is the time and  $x$  is the longitudinal coordinate. In the above system (3.1),  $\rho = \rho(x, t)$  is the density,  $A = A(x, t)$  is the cross sectional area,  $U = U(x, t)$  is the velocity averaged over the section,  $p = p(x, t)$  is the pressure,  $R = R(x, t) = \sqrt{A/\pi}$  is the tube radius, while  $\tau_s = \tau_s(x, t)$  and  $\tau_u = \tau_u(x, t)$  are the steady and the unsteady wall shear stress, respectively. Since there are six unknowns in the system (3.1), four closure relations need to be introduced. The first one is the equation of state for barotropic fluids that connects the density with the pressure:

$$\rho(p) = \rho_0 + \frac{p - p_v}{c_0^2}. \quad (3.2)$$

In (3.2) we introduced three physical parameters that are assumed to be known and constant: the reference density  $\rho_0$ , the vapor pressure  $p_v$  and the speed of sound  $c_0$ . Next, the closure for the tube law is introduced. In this case the radius is related to the pressure and here we use two simple algebraic elastic ring models. One possible choice that is very widespread in biomedical engineering is the so-called Laplace law

$$R(p) = R_0 + \frac{p - p_0}{\beta}, \quad (3.3)$$

while another closure used in mechanical applications is Hooke's law (see [104, 105]), which reads

$$R(p) = R_0 \sqrt{1 + \frac{2Wp}{E_\infty}}. \quad (3.4)$$

In equations (3.3) and (3.4)  $R_0$  is the equilibrium radius,  $p_0$  is the equilibrium pressure,  $\beta$  and  $E_\infty$  are stiffness parameters and  $W$  is a geometry factor. The *steady* wall shear stress is given by the classical Darcy-Weisbach law for laminar flow (see [91])

$$\tau_s = \lambda \rho \frac{U^2}{8}, \quad \text{with} \quad \lambda = \frac{64}{Re} \quad \text{and} \quad Re = \frac{2\rho UR}{\mu}, \quad (3.5)$$

where  $Re$  is the Reynolds number and  $\mu$  is the dynamic viscosity. The last closure regards the *unsteady* wall shear stress that is given by the Zielke convolution integral (see [162]) and which reads

$$\tau_u(x, t) = \frac{2\mu}{R} \int_0^t W'(t - t^*) \frac{\partial U(x, t^*)}{\partial t} dt^*, \quad (3.6)$$

where  $W'$  is a weighting function. More details about the efficient numerical evaluation of Eq. (3.6) will be given later.

#### 3.1.2 Staggered grids

The computational domain  $\Omega = [x_L, x_R]$  is divided in two overlapping grids. The first one is the main grid and is composed by  $N_x$  equally spaced elements of size  $\Delta x = L/N_x = \frac{x_R - x_L}{N_x}$ . The elements of this grid are indicated by the following notation  $T_i = [x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}]$ . In addition, a dual mesh is defined and it is composed by  $N_x + 1$  cells denoted by  $T_{i+\frac{1}{2}} = [x_i, x_{i+1}]$ . Note that, except for the two elements at the domain boundaries, the edges of a staggered control volume correspond to the barycenters of two consecutive elements of the main grid and vice versa (see Fig. 3.1).

#### 3.1.3 Basis functions

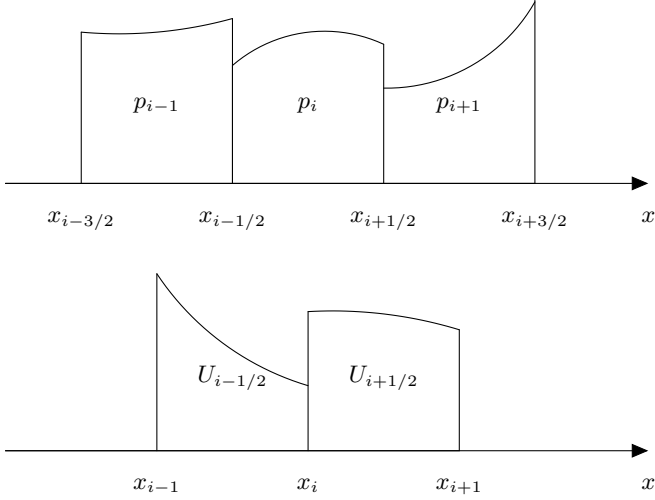
For a variable defined on the main grid, such as the pressure, the numerical solution is represented by piecewise polynomials of degree  $P$  and reads

$$p_i(x, t) = \sum_{l=1}^{P+1} \phi_l(x) \hat{p}_{i,l}(t) := \boldsymbol{\phi}(\mathbf{x}) \cdot \hat{\mathbf{p}}_i(t), \quad (3.7)$$

while for a quantity declared on the staggered grid, for example the velocity, we have

$$U_{i+\frac{1}{2}}(x, t) = \sum_{l=1}^{P+1} \psi_l(x) \hat{U}_{i+\frac{1}{2},l}(t) := \boldsymbol{\psi}(\mathbf{x}) \cdot \hat{\mathbf{U}}_{i+\frac{1}{2}}(t). \quad (3.8)$$

In equations (3.7) and (3.8)  $\hat{\mathbf{p}}_i$  and  $\hat{\mathbf{U}}_{i+\frac{1}{2}}$  are the degrees of freedom, while  $\boldsymbol{\phi}(\mathbf{x})$  and  $\boldsymbol{\psi}(\mathbf{x})$  are the so-called basis functions. Both can be mapped on an elementary control volume where a reference



**Figure 3.1:** Staggered control volumes for the DG scheme for the 1D model. Main grid used for the pressure (top) and staggered dual mesh for the velocity (bottom).

basis function  $\varphi(\xi)$  is defined:

$$\begin{aligned} \phi(x) &= \varphi(\xi) \quad \text{with} \quad x = x_i + \xi\Delta x, \quad 0 \leq \xi \leq 1 \\ \psi(x) &= \varphi(\xi) \quad \text{with} \quad x = x_{i+\frac{1}{2}} + \xi\Delta x, \quad 0 \leq \xi \leq 1. \end{aligned}$$

In this framework the reference basis functions are the Lagrange interpolation polynomials passing through the Gauss-Legendre quadrature points (see [25]) on the unit interval  $[0, 1]$ . It is obvious that the chosen nodal basis is *orthogonal* on each grid.

### 3.1.4 DG scheme for the 1D model

First, the continuity equation is multiplied by the vector of test functions  $\phi$  and is subsequently integrated over an element of the main grid:

$$\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \phi \left( \frac{\partial}{\partial t} \rho A + \frac{\partial}{\partial x} Q \right) dx = 0. \quad (3.9)$$

The quantity  $\rho A(x, t)$  is discretized as  $\phi(x) \cdot \widehat{\rho A}_i(t)$ , where the degrees of freedom are simply computed as  $\widehat{\rho A}_{i,l}(t) = \hat{\rho}_{i,l}(t) \cdot \hat{A}_{i,l}(t)$ , thanks to the use of a nodal basis. Furthermore, we introduce the quantity  $Q = \rho AU$ , which is the mass flow defined on the dual grid:  $Q_{i+\frac{1}{2}}(x, t) = \psi(x) \cdot \hat{Q}_{i+\frac{1}{2}}(t)$ . Also here the degrees of freedom of the new variable are simply given by the product  $\hat{Q}_{i+\frac{1}{2},l}(t) = \hat{\rho}_{i+\frac{1}{2},l}(t) \cdot \hat{A}_{i+\frac{1}{2},l}(t) \cdot \hat{U}_{i+\frac{1}{2},l}(t)$ .

Then, the momentum equation is multiplied by the vector of test functions  $\psi$  and integrated over a dual cell:

$$\int_{x_i}^{x_{i+1}} \psi \left( \frac{\partial Q}{\partial t} + \frac{\partial U Q}{\partial x} + A \frac{\partial p}{\partial x} \right) dx = \int_{x_i}^{x_{i+1}} \psi (-\gamma Q - \tau) dx. \quad (3.10)$$

Eq. (3.10) is written with respect to the variable  $Q$  and a friction coefficient  $\gamma = \lambda U / (4R)$  has been introduced and is discretized as  $\gamma(x, t) = \psi(x) \cdot \hat{\gamma}_{i+\frac{1}{2}}(t)$ . Moreover, in order to keep a compact notation the unsteady friction term is rewritten with respect to the new parameter  $\tau = 2\pi R \tau_u$ , defined on the dual grid as  $\tau_{i+\frac{1}{2}}(x, t) = \psi(x) \cdot \hat{\tau}_{i+\frac{1}{2}}(t)$ .

In Eq. (3.9), it is possible to discretize the time derivative while the second term is integrated by parts and, in order to have a second order approximation in time, the  $\theta$ -method is applied to the degrees of freedom of  $Q$ :

$$\begin{aligned} & \left( \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \phi \phi dx \right) \cdot \left( \frac{\widehat{\rho A}_i^{n+1} - \widehat{\rho A}_i^n}{\Delta t} + \phi(x_{i+\frac{1}{2}}^-) \psi(x_{i+\frac{1}{2}}) \cdot \hat{Q}_{i+\frac{1}{2}}^{n+\theta} \right. \\ & \left. - \phi(x_{i-\frac{1}{2}}^+) \psi(x_{i-\frac{1}{2}}) \cdot \hat{Q}_{i-\frac{1}{2}}^{n+\theta} - \left( \int_{x_{i-\frac{1}{2}}}^{x_i} \frac{\partial \phi}{\partial x} \psi dx \right) \cdot \hat{Q}_{i-\frac{1}{2}}^{n+\theta} - \left( \int_{x_i}^{x_{i+\frac{1}{2}}^-} \frac{\partial \phi}{\partial x} \psi dx \right) \cdot \hat{Q}_{i+\frac{1}{2}}^{n+\theta} \right) = 0, \end{aligned} \quad (3.11)$$

where we use the notation  $\phi \phi := \phi_k \phi_l$ ,  $\widehat{\rho A}_i^n = \widehat{\rho A}_{i,l}^n = \rho(\hat{p}_{i,l}^n) A(\hat{p}_{i,l}^{n+1})$  and  $\hat{Q}_{i+\frac{1}{2}}^{n+\theta} = (1-\theta)\hat{Q}_{i+\frac{1}{2}}^n + \theta\hat{Q}_{i+\frac{1}{2}}^{n+1}$ .

For the momentum equation, the time derivative is discretized introducing the term  $\widehat{FQ}_{i+\frac{1}{2}}^n$ , which is an explicit DG discretization of the nonlinear convective terms that will be discussed later. Then, following the approach used in [54, 70, 136, 137], the pressure term is composed by a

### 3 Staggered SI DG for axially symmetric viscous compressible flows in elastic tubes

non-conservative product which is integrated using the  $\theta$ -method and splitting the smooth contribution from the jump term. Moreover, in order to have an efficient and stable method,  $A$  is considered on the staggered grid and is evaluated at the time  $t^n$ . For the steady friction contribution, the degrees of freedom of  $\gamma$  are taken at time  $t^n$  but those of the mass flow  $Q$  are evaluated at  $t^{n+1}$ . This is done for stability reasons according to the procedure used in [91, 141], in order to deal also with potentially stiff friction terms. An explicit approach is used for the degrees of freedom of the unsteady wall shear stress, whose approximation is rather complex. In this model it is calculated using the approximation of Eq. (3.6) proposed recently in [91]:

$$\tau_u^n = \sum_{j=1}^{N_w} \tau_{u,j}^n = \sum_{j=1}^{N_w} \frac{\tau_{u,j}^{n-1} + \frac{2\mu}{R} m_j (U^n - U^{n-1})}{1 + \frac{n_j \nu}{R^2} \Delta t}, \quad (3.12)$$

where  $m_j$  and  $n_j$  are the coefficients proposed by Urbanowicz Zarzycki in [165] in order to use an approximate weighting function  $W'$  (see Eq. (3.6)) of the form:

$$W(\hat{t}) = \sum_{j=1}^{N_w} m_j e^{-n_j \hat{t}} \quad \text{with} \quad \hat{t} = \nu t / R^2. \quad (3.13)$$

The discrete momentum equation then reads

$$\begin{aligned} & \left( \int_{x_i}^{x_{i+1}} \psi \psi dx \right) \cdot \left( \frac{\hat{Q}_{i+\frac{1}{2}}^{n+1} - \widehat{FQ}_{i+\frac{1}{2}}^n}{\Delta t} \right) = \\ & - \psi(x_{i+\frac{1}{2}}) \psi(x_{i+\frac{1}{2}}) \left( \phi(x_{i+\frac{1}{2}}^+) \cdot \hat{A}_{i+\frac{1}{2}}^n \hat{p}_{i+1}^{n+\theta} - \phi(x_{i+\frac{1}{2}}^-) \cdot \hat{A}_{i+\frac{1}{2}}^n \hat{p}_i^{n+\theta} \right) \\ & - \left( \int_{x_i}^{x_{i+\frac{1}{2}}^-} \psi \psi \frac{\partial \phi}{\partial x} dx \right) \cdot \hat{A}_{i+\frac{1}{2}}^n \hat{p}_i^{n+\theta} - \left( \int_{x_{i+\frac{1}{2}}^+}^{x_{i+1}} \psi \psi \frac{\partial \phi}{\partial x} dx \right) \cdot \hat{A}_{i+\frac{1}{2}}^n \hat{p}_{i+1}^{n+\theta} \\ & - \left( \int_{x_i}^{x_{i+1}} \phi \phi dx \right) \hat{\gamma}_{i+\frac{1}{2}}^n \cdot \hat{Q}_{i+\frac{1}{2}}^{n+1} - \left( \int_{x_i}^{x_{i+1}} \phi \phi dx \right) \cdot \hat{\tau}_{i+\frac{1}{2}}^n. \end{aligned} \quad (3.14)$$

Now it is possible to introduce the following universal tensors

$$\begin{aligned} \mathbf{M}_e &= \int_0^1 \varphi(\xi) \varphi(\xi) d\xi = \int_0^1 \varphi_k(\xi) \varphi_l(\xi) d\xi = M_{e,kl}, \\ \mathbf{M} &= \int_0^1 \varphi(\xi) \varphi(\xi) \varphi(\xi) d\xi = \int_0^1 \varphi_k(\xi) \varphi_l(\xi) \varphi_m(\xi) d\xi = M_{klm}, \end{aligned} \quad (3.15)$$

$$\begin{aligned} \mathbf{R}_q &= \varphi(1)\varphi\left(\frac{1}{2}\right) - \int_{\frac{1}{2}}^1 \varphi'(\xi)\varphi\left(\xi - \frac{1}{2}\right) d\xi = \\ \varphi_k(1)\varphi_l\left(\frac{1}{2}\right) - \int_{\frac{1}{2}}^1 \varphi'_k(\xi)\varphi_l\left(\xi - \frac{1}{2}\right) d\xi &= R_{q,kl}, \end{aligned} \quad (3.16)$$

$$\begin{aligned} \mathbf{L}_q &= \varphi(0)\varphi\left(\frac{1}{2}\right) + \int_0^{\frac{1}{2}} \varphi'(\xi)\varphi\left(\xi + \frac{1}{2}\right) d\xi = \\ \varphi_k(0)\varphi_l\left(\frac{1}{2}\right) + \int_0^{\frac{1}{2}} \varphi'_k(\xi)\varphi_l\left(\xi + \frac{1}{2}\right) d\xi &= L_{q,kl}, \end{aligned} \quad (3.17)$$

$$\begin{aligned} \mathbf{R}_p &= \varphi\left(\frac{1}{2}\right)\varphi\left(\frac{1}{2}\right)\varphi(0) + \int_{\frac{1}{2}}^1 \varphi(\xi)\varphi(\xi)\varphi'\left(\xi - \frac{1}{2}\right) d\xi = \\ \varphi_k\left(\frac{1}{2}\right)\varphi_l\left(\frac{1}{2}\right)\varphi_m(0) + \int_{\frac{1}{2}}^1 \varphi_k(\xi)\varphi_l(\xi)\varphi'_m\left(\xi - \frac{1}{2}\right) d\xi &= R_{p,klm}, \end{aligned} \quad (3.18)$$

$$\begin{aligned} \mathbf{L}_p &= \varphi\left(\frac{1}{2}\right)\varphi\left(\frac{1}{2}\right)\varphi(1) - \int_0^{\frac{1}{2}} \varphi(\xi)\varphi(\xi)\varphi'\left(\xi + \frac{1}{2}\right) d\xi = \\ \varphi_k\left(\frac{1}{2}\right)\varphi_l\left(\frac{1}{2}\right)\varphi_m(1) - \int_0^{\frac{1}{2}} \varphi_k(\xi)\varphi_l(\xi)\varphi'_m\left(\xi + \frac{1}{2}\right) d\xi &= L_{p,klm}, \end{aligned} \quad (3.19)$$

in order to rewrite Eq. (3.11) and Eq. (3.14) in the more compact matrix-vector form as follows,

$$\mathbf{M}_e \cdot (\rho \mathbf{A}(\hat{\mathbf{p}}_i^{n+1}) - \rho \mathbf{A}(\hat{\mathbf{p}}_i^n)) + \frac{\Delta t}{\Delta x} (\mathbf{R}_q \cdot \hat{\mathbf{Q}}_{i+\frac{1}{2}}^{n+\theta} - \mathbf{L}_q \cdot \hat{\mathbf{Q}}_{i-\frac{1}{2}}^{n+\theta}) = 0, \quad (3.20)$$

$$\begin{aligned} \mathbf{M}_e \cdot \left( \hat{\mathbf{Q}}_{i+\frac{1}{2}}^{n+1} - \widehat{F\mathbf{Q}}_{i+\frac{1}{2}}^{n+1} \right) + \frac{\Delta t}{\Delta x} \left( \mathbf{R}_p \cdot \hat{\mathbf{A}}_{i+\frac{1}{2}}^n \hat{\mathbf{p}}_{i+1}^{n+\theta} - \mathbf{L}_p \cdot \hat{\mathbf{A}}_{i+\frac{1}{2}}^n \hat{\mathbf{p}}_i^{n+\theta} \right) = \\ - \Delta t \mathbf{M} \cdot \hat{\boldsymbol{\gamma}}_{i+\frac{1}{2}}^n \hat{\mathbf{Q}}_{i+\frac{1}{2}}^{n+1} - \Delta t \mathbf{M}_e \cdot \hat{\boldsymbol{\sigma}}_{i+\frac{1}{2}}^n, \end{aligned} \quad (3.21)$$

where the products of the rank 3 tensors with degrees of freedom are defined as  $\mathbf{R}_p \cdot \hat{\mathbf{A}}_{i+\frac{1}{2}}^n \hat{\mathbf{p}}_{i+1}^{n+\theta} = R_{p,klm} \cdot \hat{A}_{i+\frac{1}{2},l}^n \hat{p}_{i+1,m}^{n+\theta}$ ,  $\mathbf{L}_p \cdot \hat{\mathbf{A}}_{i+\frac{1}{2}}^n \hat{\mathbf{p}}_i^{n+\theta} = L_{p,klm} \cdot \hat{A}_{i+\frac{1}{2},l}^n \hat{p}_{i,m}^{n+\theta}$  and  $\mathbf{M} \cdot \hat{\gamma}_{i+\frac{1}{2}}^n \hat{\mathbf{Q}}_{i+\frac{1}{2}}^{n+1} = M_{klm} \cdot \hat{\gamma}_{i+\frac{1}{2},l}^n \hat{Q}_{i+\frac{1}{2},m}^{n+1}$ . Moreover, the matrix given by the product associated with the steady friction term is denoted by  $\mathbf{\Gamma}_{i+\frac{1}{2}}^n := \mathbf{M} \cdot \hat{\gamma}_{i+\frac{1}{2}}^n = M_{klm} \hat{\gamma}_{i+\frac{1}{2},k}^n$ . Then, the discrete momentum equation takes the form

$$\hat{\mathbf{Q}}_{i+\frac{1}{2}}^{n+1} = \left( \mathbf{M}_e + \Delta t \mathbf{\Gamma}_{i+\frac{1}{2}}^n \right)^{-1} \left( \hat{\mathbf{G}}_{i+\frac{1}{2}}^n - \theta \frac{\Delta t}{\Delta x} \left( \mathbf{R}_p \cdot \hat{\mathbf{A}}_{i+\frac{1}{2}}^n \hat{\mathbf{p}}_{i+1}^{n+1} - \mathbf{L}_p \cdot \hat{\mathbf{A}}_{i+\frac{1}{2}}^n \hat{\mathbf{p}}_i^{n+1} \right) \right), \quad (3.22)$$

where  $\hat{\mathbf{G}}_{i+\frac{1}{2}}^n$  accounts for all the known terms at the time  $t^n$ :

$$\hat{\mathbf{G}}_{i+\frac{1}{2}}^n = \mathbf{M}_e \cdot \widehat{\mathbf{F}} \hat{\mathbf{Q}}_{i+\frac{1}{2}}^n - (1-\theta) \frac{\Delta t}{\Delta x} \left( \mathbf{R}_p \cdot \hat{\mathbf{A}}_{i+\frac{1}{2}}^n \hat{\mathbf{p}}_{i+1}^n - \mathbf{L}_p \cdot \hat{\mathbf{A}}_{i+\frac{1}{2}}^n \hat{\mathbf{p}}_i^n \right) - \Delta t \mathbf{M}_e \cdot \hat{\tau}_{i+\frac{1}{2}}^n. \quad (3.23)$$

Inserting the discrete momentum equation given by Eq. (3.22) into the discrete continuity equation (3.20) yields the following system for the discrete pressure:

$$\begin{aligned} & \mathbf{M}_e \cdot \rho \mathbf{A}(\hat{\mathbf{p}}_i^{n+1}) - \theta^2 \frac{\Delta t^2}{\Delta x^2} \mathbf{R}_q [\mathbf{M}_e + \Delta t \mathbf{\Gamma}_{i+\frac{1}{2}}^n]^{-1} \mathbf{R}_p \cdot \hat{\mathbf{A}}_{i+\frac{1}{2}}^n \hat{\mathbf{p}}_{i+1}^{n+1} \\ & - \theta^2 \frac{\Delta t^2}{\Delta x^2} \mathbf{L}_q [\mathbf{M}_e + \Delta t \mathbf{\Gamma}_{i-\frac{1}{2}}^n]^{-1} \mathbf{L}_p \cdot \hat{\mathbf{A}}_{i-\frac{1}{2}}^n \hat{\mathbf{p}}_{i-1}^{n+1} + \\ & \theta^2 \frac{\Delta t^2}{\Delta x^2} \left( \mathbf{R}_q [\mathbf{M}_e + \Delta t \mathbf{\Gamma}_{i+\frac{1}{2}}^n]^{-1} \mathbf{L}_p \cdot \hat{\mathbf{A}}_{i+\frac{1}{2}}^n + \mathbf{L}_q [\mathbf{M}_e + \Delta t \mathbf{\Gamma}_{i-\frac{1}{2}}^n]^{-1} \mathbf{R}_p \cdot \hat{\mathbf{A}}_{i-\frac{1}{2}}^n \right) \cdot \hat{\mathbf{p}}_i^{n+1} \\ & = \mathbf{M}_e \cdot \rho \mathbf{A}(\hat{\mathbf{p}}_i^n) - (1-\theta) \frac{\Delta t}{\Delta x} \left( \mathbf{R}_q \cdot \hat{\mathbf{Q}}_{i+\frac{1}{2}}^n - \mathbf{L}_q \cdot \hat{\mathbf{Q}}_{i-\frac{1}{2}}^n \right) \\ & - \theta \frac{\Delta t}{\Delta x} (\mathbf{R}_q [\mathbf{M}_e + \Delta t \mathbf{\Gamma}_{i+\frac{1}{2}}^n]^{-1} \cdot \hat{\mathbf{G}}_{i+\frac{1}{2}}^n - \mathbf{L}_q [\mathbf{M}_e + \Delta t \mathbf{\Gamma}_{i-\frac{1}{2}}^n]^{-1} \cdot \hat{\mathbf{G}}_{i-\frac{1}{2}}^n), \end{aligned} \quad (3.24)$$

which takes the form of the following mildly non-linear system for the unknown pressure degrees of freedom  $\hat{\mathbf{p}}^{n+1}$ :

$$\rho \mathbf{A}(\hat{\mathbf{p}}^{n+1}) + \mathbf{T} \cdot \hat{\mathbf{p}}^{n+1} = \mathbf{b}^n. \quad (3.25)$$

In Eq. (3.25)  $\rho \mathbf{A}(\hat{\mathbf{p}}^{n+1})$  is the mildly non-linear contribution that affects only the diagonal of the system (note that our chosen nodal basis is *orthogonal*),  $\mathbf{b}^n$  is the known right hand side and  $\mathbf{T}$  is a symmetric block three-diagonal matrix. The degrees of freedom of the pressure  $\hat{\mathbf{p}}^{n+1}$  are obtained using the Newton-type algorithm for mildly nonlinear systems proposed by Casulli et al. [18, 19, 37, 38] and which has already successfully been used in [28, 34, 72, 141]. Once the pressure is known, the density and the cross sectional area on the main grid are computed using the relations (3.2), (3.3) or (3.4), while  $\hat{\mathbf{Q}}_{i+\frac{1}{2}}^{n+1}$  is directly given by the discrete momentum equation (3.22). The degrees of freedom of the area on the dual grid are obtained using the following  $L_2$  averaging operator

$$\hat{\mathbf{A}}_{i+\frac{1}{2}}^n = \mathbf{M}_e^{-1} \left( \mathbf{M}_L \cdot \hat{\mathbf{A}}_i^n + \mathbf{M}_R \cdot \hat{\mathbf{A}}_{i+1}^n \right), \quad (3.26)$$



where  $\mathbf{M}_{\mathbf{L}}$  and  $\mathbf{M}_{\mathbf{R}}$  are universal matrices defined as follows

$$\begin{aligned}\mathbf{M}_{\mathbf{L}} &= \int_0^{\frac{1}{2}} \varphi(\xi) \varphi(\xi + \frac{1}{2}) d\xi = \int_0^{\frac{1}{2}} \varphi_k(\xi) \varphi_l(\xi + \frac{1}{2}) d\xi = M_{L,kl} \\ \mathbf{M}_{\mathbf{R}} &= \int_{\frac{1}{2}}^1 \varphi(\xi) \varphi(\xi - \frac{1}{2}) d\xi = \int_{\frac{1}{2}}^1 \varphi_k(\xi) \varphi_l(\xi - \frac{1}{2}) d\xi = M_{R,kl}.\end{aligned}\tag{3.27}$$

Then, thanks to the use of a nodal basis, the degrees of freedom of the velocity can be computed pointwise simply as

$$\hat{U}_{i+\frac{1}{2},l}^{n+1} = \frac{\hat{Q}_{i+\frac{1}{2},l}^{n+1}}{\hat{\rho}_{i+\frac{1}{2},l}^{n+1} \hat{A}_{i+\frac{1}{2},l}^{n+1}}.\tag{3.28}$$

### 3.1.5 Non-linear convective terms for the 1D model

In order to complete the description of the numerical scheme, the procedure to compute the term  $\widehat{\mathbf{FQ}}_i^n$  is given. At first, the velocity and the mass flow are projected on the main longitudinal grid applying the  $L_2$  average in Eq. (3.26) properly applied to the quantities involved. Then, for stability reasons,  $\widehat{\mathbf{FQ}}_i^n$  is computed using a Runge Kutta Discontinuous Galerkin scheme based on the classical third order TVD Runge-Kutta method:

$$\begin{aligned}\hat{\mathbf{k}}_1 &= \hat{\mathbf{Q}}^n + \Delta t \mathbf{L}_h \left( \hat{\mathbf{U}}^n, \hat{\mathbf{Q}}^n \right), \\ \hat{\mathbf{k}}_2 &= \frac{3}{4} \hat{\mathbf{Q}}^n + \frac{1}{4} \hat{\mathbf{k}}_1 + \frac{1}{4} \Delta t \mathbf{L}_h \left( \hat{\mathbf{U}}^n, \hat{\mathbf{k}}_1 \right), \\ \widehat{\mathbf{FQ}}^n &= \frac{1}{3} \hat{\mathbf{Q}}^n + \frac{2}{3} \hat{\mathbf{k}}_2 + \frac{2}{3} \Delta t \mathbf{L}_h \left( \mathbf{U}^n, \hat{\mathbf{k}}_2 \right),\end{aligned}\tag{3.29}$$

where the operator  $\mathbf{L}_h \left( \hat{\mathbf{U}}, \hat{\mathbf{Q}} \right)$  is computed for each cell  $i$  as follows

$$\mathbf{L}_h \left( \hat{\mathbf{U}}, \hat{\mathbf{Q}} \right)_i = -\frac{1}{\Delta x} \mathbf{Me}^{-1} \left( \varphi(1) f_{i+\frac{1}{2}} - \varphi(0) f_{i-\frac{1}{2}} - \int_0^1 \varphi'(\xi) \varphi(\xi) d\xi \cdot \hat{\mathbf{f}}_i \right).\tag{3.30}$$

Since the physical flux of the nonlinear convective terms is  $f = UQ$ , in Eq. (4.18) we have that  $\hat{\mathbf{f}}_i = \hat{\mathbf{U}}_i \hat{\mathbf{Q}}_i = \hat{U}_{i,l} \hat{Q}_{i,l}$ , while the numerical flux  $f_{i+\frac{1}{2}}$  is a function of the two boundary extrapolated values at the interface  $x_{i+\frac{1}{2}}$ , which are given by  $U_{i+\frac{1}{2}}^- = \varphi(\mathbf{1}) \cdot \hat{\mathbf{U}}_i$  and  $U_{i+\frac{1}{2}}^+ = \varphi(\mathbf{0}) \cdot \hat{\mathbf{U}}_{i+1}$ . Here we use the simple Rusanov flux, which reads

$$f_{i+\frac{1}{2}} = \frac{1}{2} \left( U_{i+\frac{1}{2}}^+ Q_{i+\frac{1}{2}}^+ + U_{i+\frac{1}{2}}^- Q_{i+\frac{1}{2}}^- \right) - \frac{1}{2} \max \left( |U_{i+\frac{1}{2}}^+|, |U_{i+\frac{1}{2}}^-| \right) \left( Q_{i+\frac{1}{2}}^+ - Q_{i+\frac{1}{2}}^- \right).\tag{3.31}$$

At the end, the non-linear convective terms are projected back to the staggered dual mesh using again the  $L_2$  average. If the evaluation of the non-linear terms can be neglected one can simply

set  $\widehat{\mathbf{FQ}}_{i+\frac{1}{2}}^n = \hat{\mathbf{Q}}_{i+\frac{1}{2}}^n$  and consequently the scheme becomes unconditionally stable. However, the time step  $\Delta t$  should be chosen small enough to reduce numerical viscosity and to resolve transient features properly. On the contrary, when the convective terms have to be calculated then the admissible time step  $\Delta t$  is given by the usual CFL condition for RKDG schemes, which in this case reads  $\Delta t < \text{CFL} \frac{\Delta x}{\max(|U|)}$  with  $\text{CFL} < (2P + 1)^{-1}$ .

## 3.2 Staggered semi-implicit DG scheme applied to the $2D_{xr}$ model

### 3.2.1 Governing equations of the $2D_{xr}$ model

In some situations, for example when the accurate knowledge of the velocity profile is important, one may be interested in a more complete model. Consequently, when the radial scale is much smaller than the longitudinal one, it is still reasonable to assume axial symmetry and hydrostatic equilibrium, which means that pressure and density are constant within the cross sectional area. Moreover, the contribution of the viscosity in longitudinal direction can be neglected and, at the end, the Navier-Stokes equations reduce to the following system of PDEs called  $2D_{xr}$  model (see [56, 91]) composed by the continuity equation and by the momentum equation in axial direction:

$$\frac{\partial \rho r}{\partial t} + \frac{\partial}{\partial x}(\rho u r) + \frac{\partial}{\partial r}(\rho r w) = 0, \quad (3.32a)$$

$$\rho \left( \frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} + w \frac{\partial u}{\partial r} \right) + \frac{\partial p}{\partial x} = \frac{\mu}{r} \frac{\partial}{\partial r} \left( r \frac{\partial u}{\partial r} \right). \quad (3.32b)$$

In (3.32)  $u = u(x, r, t)$  and  $w = w(x, r, t)$  are the velocity components in  $x$  and  $r$  direction, respectively. The no-slip condition at tube wall ( $r = R$ ) and the symmetry condition at the center of the pipe ( $r = 0$ ) state:

$$u(x, R, t) = 0 \quad \text{and} \quad \frac{\partial u(x, 0, t)}{\partial r} = 0. \quad (3.33)$$

Then, the kinematic boundary condition at the moving wall boundary reads (see [34, 72])

$$\frac{\partial R}{\partial t} + u \frac{\partial R}{\partial x} - w = 0 \quad \text{for} \quad r = R, \quad (3.34)$$

and integration over the cross sectional area of Eq. (3.32a) coupled with Eq. (3.34) (see [56, 91]) yields the following integral continuity equation

$$\frac{\partial \rho A}{\partial t} + 2\pi \frac{\partial}{\partial x} \int_0^R \rho r u dr = 0. \quad (3.35)$$

### 3.2.2 Computational grids and basis functions

The computational domain defined by  $\Omega(t) = \{(x, r) : x \in [x_L, x_R] \cup r \in [0, R(x, t)]\}$  in general is non constant in time. Similar to the numerical method described in Section 3.1, a main grid and a dual grid are introduced in  $x$  direction formed by  $N_x$  and by  $N_x + 1$  cells, respectively.

Due to the assumption of hydrostatic equilibrium, the pressure, the density and the area are variables that depend only on  $x$  and  $t$ . Consequently, these quantities are discretized exactly as it has been done in Section 3.1, however here we use the apex  $x$  for the basis functions so that they can be distinguished from the basis functions in  $r$  direction that will be discussed later:

$$p_i(x, t) = \sum_{l=1}^{P^x+1} \phi_l^x(x) \hat{p}_{i,l}(t) = \phi^x(x) \cdot \hat{\mathbf{p}}_i(t). \quad (3.36)$$

The components of the velocity  $u$  and  $w$  depend also on the coordinate  $r$  and consequently two staggered meshes are introduced in every degree of freedom of the longitudinal meshes. The main radial grid is composed by  $N_r$  rings while for the dual radial grid there are  $N_r + 1$  rings that are bounded by the barycentres of the elements of the main radial grid. Then, the basis functions for the radial meshes are introduced. The polynomial basis functions  $\phi^r(r)$  and  $\psi^r(r)$  are characterized by a polynomial degree  $P^r$ . Similar to the one-dimensional model, they are computed with respect the function  $\varphi^r$  on a reference domain using the Lagrange interpolation polynomials passing through the Gauss-Legendre quadrature nodes. The adoption of possibly different polynomial degrees for the basis functions in  $x$  and  $r$  is justified by the fact that  $u$  and  $w$  change differently along the two coordinates. The longitudinal velocity  $u$  is defined on a control volume  $\Omega_{i+\frac{1}{2},k}^n = [x_i, x_{i+1}] \times [r_{k-\frac{1}{2}}^n(x), r_{k+\frac{1}{2}}^n(x)]$  that belongs to the staggered longitudinal grid and to the main radial grid. In terms of basis functions, the discrete velocity  $u$  is written as follows:

$$u_{i+\frac{1}{2},k}(x, r, t) = \sum_l^{P^x+1} \sum_m^{P^r+1} \psi_l^x(x) \phi_m^r(r) \hat{u}_{i+\frac{1}{2},k,l,m}(t) = \psi^x(x) \phi^r(r) \cdot \hat{\mathbf{u}}_{i+\frac{1}{2},k}(t). \quad (3.37)$$

The radial velocity  $w$  is declared on the main longitudinal mesh and on the staggered radial mesh in the control volume  $\Omega_{i,k+\frac{1}{2}}^n = [x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}] \times [r_k^n(x), r_{k+1}^n(x)]$ . Numerically it is expressed as

$$w_{i,k+\frac{1}{2}}(x, r, t) = \sum_l^{P^x+1} \sum_m^{P^r+1} \phi_l^x(x) \psi_m^r(r) \hat{w}_{i,k+\frac{1}{2},l,m}(t) = \phi^x(x) \psi^r(r) \cdot \hat{\mathbf{w}}_{i,k+\frac{1}{2}}(t). \quad (3.38)$$

### 3.2.3 DG scheme for the $2D_{xr}$ model - Integral continuity equation

It is easy to observe that the integral  $2\pi \int_0^R \rho u r dr$  is the total cross sectional mass flow  $Q$  used also in the 1D model. Consequently, the PDE Eq. (3.35) corresponds to the PDE in Eq. (3.1a) and also its discretization is carried out in the same way done in the previous section. However, since for this model the basis functions in  $x$  direction and  $r$  direction are characterized by different polynomial degrees, the universal tensors introduced in Section 3.1 now are denoted by the apex  $\mathbf{x}$ . This is done because in the next subsection other tensors will be declared and they will differ by the chosen set of basis function. The discrete integral continuity equation thus can be written as in the 1D model and reads

$$\mathbf{M}_e^{\mathbf{x}} \cdot (\rho \mathbf{A}(\hat{\mathbf{p}}_i^{n+1}) - \rho \mathbf{A}(\hat{\mathbf{p}}_i^n)) + \frac{\Delta t}{\Delta x} (\mathbf{R}_q^{\mathbf{x}} \cdot \hat{\mathbf{Q}}_{i+\frac{1}{2}}^{n+\theta} - \mathbf{L}_q^{\mathbf{x}} \cdot \hat{\mathbf{Q}}_{i-\frac{1}{2}}^{n+\theta}) = 0. \quad (3.39)$$

In the next sections we will discuss in detail how to compute the degrees of freedom  $\hat{\mathbf{Q}}_{i+\frac{1}{2}}^{n+\theta}$  of the averaged mass flow from the velocity components defined on the radial grid, taking into account also the viscous terms.

### 3.2.4 DG scheme for the $2D_{xr}$ model - Momentum equation - Part A

In order to make the complete method understandable, the integration of the momentum equation is first developed such that a low order approach is adopted in  $x$  direction. Hence, the polynomial degree  $P^x$  of the basis functions  $\phi^x$  is 0 and consequently  $\phi^x(x) = 1$ . The momentum equation is multiplied by a test function that depends only on  $r$  and then it is integrated over a control volume  $\Omega_{i+\frac{1}{2},k}^n$

$$\begin{aligned} & \int_{x_i}^{x_{i+1}} \int_{r_{i+\frac{1}{2},k-\frac{1}{2}}^n}^{r_{i+\frac{1}{2},k+\frac{1}{2}}^n} \phi^{\mathbf{r}} 2\pi\rho r \left( \frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} + w \frac{\partial u}{\partial r} \right) dr dx = \\ & \int_{x_i}^{x_{i+1}} \int_{r_{i+\frac{1}{2},k-\frac{1}{2}}^n}^{r_{i+\frac{1}{2},k+\frac{1}{2}}^n} \phi^{\mathbf{r}} \left( -r \frac{\partial p}{\partial x} + \mu \frac{\partial}{\partial r} f^r \right) dr dx, \end{aligned} \quad (3.40)$$

where  $f^r$  is the viscous flux in radial direction which is equal to  $r \frac{\partial u}{\partial r}$  and it is declared on the dual radial grid  $f_{i+1/2,k+1/2}^r = \phi^{\mathbf{r}} \cdot \hat{\mathbf{f}}_{i+1/2,k+1/2}^r$ . Then, a new variable  $q = 2\pi\rho ur$  is introduced and is discretized as  $q_{i+1/2,k} = \phi^{\mathbf{r}} \cdot \hat{\mathbf{q}}_{i+1/2,k}$ . On the left hand side of the momentum equation the nonlinear convective terms are again discretized introducing the operator  $\widehat{\mathbf{F}} \hat{\mathbf{q}}_{i+1/2,k}$ . At the right hand side of Eq. (3.40) the pressure gradient is discretized using the  $\theta$ -method and the viscous term is integrated by parts, which yields

$$\begin{aligned} & \int_{r_{i+\frac{1}{2},k-\frac{1}{2}}^n}^{r_{i+\frac{1}{2},k+\frac{1}{2}}^n} \phi^{\mathbf{r}} \phi^{\mathbf{r}} dr \cdot \left( \frac{\hat{\mathbf{q}}_{i+\frac{1}{2},k}^{n+1} - \widehat{\mathbf{F}} \hat{\mathbf{q}}_{i+\frac{1}{2},k}^n}{\Delta t} \right) = - \int_{r_{i+\frac{1}{2},k-\frac{1}{2}}^n}^{r_{i+\frac{1}{2},k+\frac{1}{2}}^n} 2\pi r \phi^{\mathbf{r}} dr \cdot \left( \frac{\hat{p}_{i+1}^{n+\theta} - \hat{p}_i^{n+\theta}}{\Delta x} \right) \\ & + 2\pi\mu [\phi^{\mathbf{r}}(r_{i+\frac{1}{2},k+\frac{1}{2}}^{n-}) \psi^{\mathbf{r}}(r_{i+\frac{1}{2},k+\frac{1}{2}}^n) \cdot \hat{\mathbf{f}}_{i+\frac{1}{2},k+\frac{1}{2}}^{n+\theta} \\ & - \phi^{\mathbf{r}}(r_{i+\frac{1}{2},k-\frac{1}{2}}^{n+}) \psi^{\mathbf{r}}(r_{i+\frac{1}{2},k-\frac{1}{2}}^n) \cdot \hat{\mathbf{f}}_{i+\frac{1}{2},k-\frac{1}{2}}^{n+\theta} \\ & - \int_{r_{i+\frac{1}{2},k-\frac{1}{2}}^{n+}}^{r_{i+\frac{1}{2},k}^n} \frac{\partial \phi^{\mathbf{r}}}{\partial r} \psi^{\mathbf{r}} dr \cdot \hat{\mathbf{f}}_{i+\frac{1}{2},k-\frac{1}{2}}^{n+\theta} - \int_{r_{i+\frac{1}{2},k}^n}^{r_{i+\frac{1}{2},k+\frac{1}{2}}^{n-}} \frac{\partial \phi^{\mathbf{r}}}{\partial r} \psi^{\mathbf{r}} dr \cdot \hat{\mathbf{f}}_{i+\frac{1}{2},k+\frac{1}{2}}^{n+\theta}]. \end{aligned} \quad (3.41)$$

The viscous flux on the staggered control volumes in  $r$  direction is obtained by considering the smooth parts and the jump terms arising in the radial velocity profile of  $u$  as follows:

$$\begin{aligned}
 & \int_{r_{i+\frac{1}{2},k}^n}^{r_{i+\frac{1}{2},k+1}^n} \psi^{\mathbf{r}} \psi^{\mathbf{r}} dr \cdot \hat{\mathbf{f}}_{i+\frac{1}{2},k+\frac{1}{2}}^{n+\theta} = \int_{r_{i+\frac{1}{2},k}^n}^{r_{i+\frac{1}{2},k+1}^n} \psi^{\mathbf{r}} r \frac{\partial u}{\partial r} dr = \\
 & \int_{r_{i+\frac{1}{2},k}^n}^{r_{i+\frac{1}{2},k+\frac{1}{2}}^n} \psi^{\mathbf{r}} r \frac{\partial \phi^{\mathbf{r}}}{\partial r} dr \cdot \hat{\mathbf{u}}_{i+\frac{1}{2},k}^{n+\theta} + \int_{r_{i+\frac{1}{2},k+\frac{1}{2}}^n}^{r_{i+\frac{1}{2},k+1}^n} \psi^{\mathbf{r}} r \frac{\partial \phi^{\mathbf{r}}}{\partial r} dr \cdot \hat{\mathbf{u}}_{i+\frac{1}{2},k+1}^{n+\theta} \\
 & + r_{i+\frac{1}{2},k+\frac{1}{2}}^n \psi^{\mathbf{r}}(r_{i+\frac{1}{2},k+\frac{1}{2}}^n) \left( \phi^{\mathbf{r}}(r_{i+\frac{1}{2},k+\frac{1}{2}}^n) \cdot \hat{\mathbf{u}}_{i+\frac{1}{2},k+1}^{n+\theta} - \phi^{\mathbf{r}}(r_{i+\frac{1}{2},k+\frac{1}{2}}^n) \cdot \hat{\mathbf{u}}_{i+\frac{1}{2},k}^{n+\theta} \right). \quad (3.42)
 \end{aligned}$$

Now, introducing the following tensors,

$$\mathbf{M}_{\mathbf{e}}^{\mathbf{r}} = \int_0^1 \varphi^{\mathbf{r}}(\xi) \varphi^{\mathbf{r}}(\xi) d\xi \quad (3.43)$$

$$\mathbf{A}_{\mathbf{r}}^n_{i+\frac{1}{2},k} = 2\pi \left( r_{i+\frac{1}{2},k-\frac{1}{2}}^n \int_0^1 \varphi^{\mathbf{r}}(\xi) (1-\xi) d\xi + r_{i+\frac{1}{2},k+\frac{1}{2}}^n \int_0^1 \varphi^{\mathbf{r}}(\xi) \xi d\xi \right),$$

$$\begin{aligned}
 \mathbf{R}_{\mathbf{u}}^{\mathbf{r}} &= \varphi^{\mathbf{r}}(1) \varphi^{\mathbf{r}}\left(\frac{1}{2}\right) - \int_{\frac{1}{2}}^1 \varphi^{\mathbf{r}}(\xi)' \varphi^{\mathbf{r}}\left(\xi - \frac{1}{2}\right) d\xi, \\
 \mathbf{L}_{\mathbf{u}}^{\mathbf{r}} &= \varphi^{\mathbf{r}}(0) \varphi^{\mathbf{r}}\left(\frac{1}{2}\right) + \int_0^{\frac{1}{2}} \varphi^{\mathbf{r}}(\xi)' \varphi^{\mathbf{r}}\left(\xi + \frac{1}{2}\right) d\xi,
 \end{aligned} \quad (3.44)$$

the integrated momentum equation in Eq. (3.41) is rewritten in the compact matrix vector form

$$\begin{aligned}
 \mathbf{M}_{\mathbf{e}}^{\mathbf{r}} \left( \frac{\hat{\mathbf{q}}_{i+\frac{1}{2},k}^{n+1} - \widehat{\mathbf{F}} \mathbf{q}_{i+\frac{1}{2},k}^n}{\Delta t} \right) \Delta r_{i+\frac{1}{2},k}^n &= -\mathbf{A}_{\mathbf{r}}^n_{i+\frac{1}{2},k} \cdot \left( \frac{p_{i+1}^{n+\theta} - p_i^{n+\theta}}{\Delta x} \right) \Delta r_{i+\frac{1}{2},k}^n \\
 &+ \mathbf{R}_{\mathbf{u}}^{\mathbf{r}} \cdot \hat{\mathbf{f}}_{i+\frac{1}{2},k+\frac{1}{2}}^{n+\theta} - \mathbf{L}_{\mathbf{u}}^{\mathbf{r}} \cdot \hat{\mathbf{f}}_{i+\frac{1}{2},k-\frac{1}{2}}^{n+\theta}.
 \end{aligned} \quad (3.45)$$

Also the viscous flux defined in Eq. (3.42) and used in (3.45) above can be compactly written as

$$\hat{\mathbf{f}}_{i+\frac{1}{2},k+\frac{1}{2}}^{n+\theta} = \frac{(\mathbf{M}_{\mathbf{e}}^{\mathbf{r}})^{-1}}{\Delta r_{i+\frac{1}{2},k+\frac{1}{2}}^n} \cdot \left( \mathbf{F}_{i+\frac{1}{2},k+\frac{1}{2}}^+ \cdot \hat{\mathbf{u}}_{i+\frac{1}{2},k+1}^{n+\theta} - \mathbf{F}_{i+\frac{1}{2},k+\frac{1}{2}}^- \cdot \hat{\mathbf{u}}_{i+\frac{1}{2},k}^{n+\theta} \right), \quad (3.46)$$

### 3 Staggered SI DG for axially symmetric viscous compressible flows in elastic tubes

where  $\mathbf{F}_{i+\frac{1}{2},k+\frac{1}{2}}^+$  and  $\mathbf{F}_{i+\frac{1}{2},k+\frac{1}{2}}^-$  are tensors that are computed via the following expressions

$$\begin{aligned}\mathbf{F}_{i+\frac{1}{2},k+\frac{1}{2}}^+ &= - \left( r_{i+\frac{1}{2},k}^n \mathbf{F}_{\mathbf{L1}} + r_{i+\frac{1}{2},k+\frac{1}{2}}^n \mathbf{F}_{\mathbf{L2}} + r_{i+\frac{1}{2},k+1}^n \mathbf{F}_{\mathbf{L3}} \right), \\ \mathbf{F}_{i+\frac{1}{2},k+\frac{1}{2}}^- &= \left( r_{i+\frac{1}{2},k}^n \mathbf{F}_{\mathbf{R1}} + r_{i+\frac{1}{2},k+\frac{1}{2}}^n \mathbf{F}_{\mathbf{R2}} + r_{i+\frac{1}{2},k+1}^n \mathbf{F}_{\mathbf{R3}} \right),\end{aligned}\quad (3.47)$$

based on the following universal matrices

$$\begin{aligned}\mathbf{F}_{\mathbf{L1}} &= \int_0^{\frac{1}{2}} \boldsymbol{\varphi}^{\mathbf{r}}(\xi) \frac{\partial \boldsymbol{\varphi}^{\mathbf{r}}(\xi + \frac{1}{2})}{\partial \xi} (1 - \xi) d\xi, \\ \mathbf{F}_{\mathbf{L2}} &= \boldsymbol{\varphi}^{\mathbf{r}}(\frac{1}{2}) \boldsymbol{\varphi}^{\mathbf{r}}(1), \\ \mathbf{F}_{\mathbf{L3}} &= \int_0^{\frac{1}{2}} \boldsymbol{\varphi}^{\mathbf{r}}(\xi) \frac{\partial \boldsymbol{\varphi}^{\mathbf{r}}(\xi + \frac{1}{2})}{\partial \xi} \xi d\xi, \\ \mathbf{F}_{\mathbf{R1}} &= \int_{\frac{1}{2}}^1 \boldsymbol{\varphi}^{\mathbf{r}}(\xi) \frac{\partial \boldsymbol{\varphi}^{\mathbf{r}}(\xi - \frac{1}{2})}{\partial \xi} (1 - \xi) d\xi, \\ \mathbf{F}_{\mathbf{R2}} &= -\boldsymbol{\varphi}^{\mathbf{r}}(\frac{1}{2}) \boldsymbol{\varphi}^{\mathbf{r}}(0), \\ \mathbf{F}_{\mathbf{R3}} &= \int_{\frac{1}{2}}^1 \boldsymbol{\varphi}^{\mathbf{r}}(\xi) \frac{\partial \boldsymbol{\varphi}^{\mathbf{r}}(\xi - \frac{1}{2})}{\partial \xi} \xi d\xi.\end{aligned}\quad (3.48)$$

In order to have a relation between the degrees of freedom of  $q$  with those of  $u$ , we use again standard  $L_2$  projection. Hence, the expression  $q = 2\pi\rho r u$  is multiplied by the test function  $\boldsymbol{\phi}^{\mathbf{r}}$  and integrated in  $r$  direction

$$\int_{r_{i+\frac{1}{2},k-\frac{1}{2}}^n}^{r_{i+\frac{1}{2},k+\frac{1}{2}}^n} \boldsymbol{\phi}^{\mathbf{r}} \boldsymbol{\phi}^{\mathbf{r}} dr \cdot \hat{\mathbf{q}}_{i+\frac{1}{2},k}^{n+1} = 2\pi\rho_{i+\frac{1}{2}}^n \int_{r_{i+\frac{1}{2},k-\frac{1}{2}}^n}^{r_{i+\frac{1}{2},k+\frac{1}{2}}^n} \boldsymbol{\phi}^{\mathbf{r}} \boldsymbol{\phi}^{\mathbf{r}} r dr \cdot \hat{\mathbf{u}}_{i+\frac{1}{2},k}^{n+1} \quad (3.50)$$

that yields

$$\mathbf{M}_{\mathbf{e}}^{\mathbf{r}} \cdot \hat{\mathbf{q}}_{i+\frac{1}{2},k}^{n+1} = 2\pi\rho_{i+\frac{1}{2}}^n \mathbf{M}_{i+\frac{1}{2},k}^{\mathbf{r}} \cdot \hat{\mathbf{u}}_{i+\frac{1}{2},k}^{n+1}, \quad (3.51)$$

where  $\mathbf{M}_{i+\frac{1}{2},k}^{\mathbf{r}}$  is a matrix defined as:

$$\mathbf{M}_{i+\frac{1}{2},k}^{\mathbf{r}} = r_{i+\frac{1}{2},k-\frac{1}{2}}^n \int_0^1 \boldsymbol{\varphi}^{\mathbf{r}}(\xi) \boldsymbol{\varphi}^{\mathbf{r}}(\xi) (1 - \xi) d\xi + r_{i+\frac{1}{2},k+\frac{1}{2}}^n \int_0^1 \boldsymbol{\varphi}^{\mathbf{r}}(\xi) \boldsymbol{\varphi}^{\mathbf{r}}(\xi) \xi d\xi. \quad (3.52)$$

The substitution of Eq. (3.46) into the integrated momentum equation (3.45) gives the following expression

$$\begin{aligned}
 & -\theta^2 \Delta t \frac{\mathbf{R}_u^r \cdot (\mathbf{M}_e^r)^{-1} \cdot \mathbf{F}_{i+\frac{1}{2},k+\frac{1}{2}}^+ \cdot (\mathbf{M}^r_{i+\frac{1}{2},k+1})^{-1} \cdot \mathbf{M}_e^r}{\rho_{i+\frac{1}{2},k+\frac{1}{2}}^n \Delta r_{i+\frac{1}{2},k+\frac{1}{2}}^n} \cdot \hat{q}_{i+\frac{1}{2},k+1}^{n+1} + \mathbf{M}_e^r \Delta r_{i+\frac{1}{2},k} \cdot \hat{q}_{i+\frac{1}{2},k}^{n+1} \\
 & \left[ +\theta^2 \Delta t \left( \frac{\mathbf{R}_u^r \cdot (\mathbf{M}_e^r)^{-1} \cdot \mathbf{F}_{i+\frac{1}{2},k+\frac{1}{2}}^+}{\Delta r_{i+\frac{1}{2},k+\frac{1}{2}}^n} + \frac{\mathbf{L}_u^r \cdot (\mathbf{M}_e^r)^{-1} \cdot \mathbf{F}_{i+\frac{1}{2},k-\frac{1}{2}}^+}{\Delta r_{i+\frac{1}{2},k-\frac{1}{2}}^n} \right) \cdot \frac{(\mathbf{M}^r_{i+\frac{1}{2},k})^{-1} \cdot \mathbf{M}_e^r}{\rho_{i+\frac{1}{2}}^n} \right] \cdot \hat{q}_{i+\frac{1}{2},k}^{n+1} \\
 & -\theta^2 \Delta t \frac{\mathbf{L}_u^r \cdot (\mathbf{M}_e^r)^{-1} \cdot \mathbf{F}_{i+\frac{1}{2},k-\frac{1}{2}}^- \cdot (\mathbf{M}^r_{i+\frac{1}{2},k-1})^{-1} \cdot \mathbf{M}_e^r}{\rho_{i+\frac{1}{2}}^n \Delta r_{i+\frac{1}{2},k-\frac{1}{2}}^n} \cdot \hat{q}_{i+\frac{1}{2},k-1}^{n+1} = \mathbf{M}_e^r \cdot \widehat{\mathbf{F}} \hat{q}_{i+\frac{1}{2},k}^n \Delta r_{i+\frac{1}{2},k} \\
 & -\Delta r_{i+\frac{1}{2},k}^n \frac{\Delta t}{\Delta x} \mathbf{A}_r \mathbf{r}_{i+\frac{1}{2},k} (p_{i+1}^{n+\theta} - p_i^{n+\theta}) + \Delta t (1-\theta) (\mathbf{R}_u^r \cdot \hat{\mathbf{f}}_{i+\frac{1}{2},k+\frac{1}{2}}^n - \mathbf{L}_u^r \cdot \hat{\mathbf{f}}_{i+\frac{1}{2},k-\frac{1}{2}}^n)
 \end{aligned} \tag{3.53}$$

where  $\hat{q}_{i+\frac{1}{2},k}^n$  accounts for all the known terms at time  $t^n$ . Eq. (3.53) written for all  $k = 1, \dots, N_r$  forms the linear system

$$\mathbf{D}_{i+\frac{1}{2}}^n \cdot \mathbf{Q}_{i+\frac{1}{2}}^{n+1} = \mathbf{G}_{i+\frac{1}{2}}^n - \theta \frac{\Delta t}{\Delta x} \mathbf{A}_r \mathbf{r}_{i+\frac{1}{2}}^n (p_{i+1}^{n+1} - p_i^{n+1}). \tag{3.54}$$

In Eq. (3.54)  $\mathbf{D}_{i+\frac{1}{2}}^n$  is a non-symmetric block three-diagonal matrix with block size  $(P^r + 1) \times (P^r + 1)$ , while  $\mathbf{Q}_{i+\frac{1}{2}}^{n+1}$ ,  $\mathbf{G}_{i+\frac{1}{2}}^n$  and  $\mathbf{A}_r \mathbf{r}_{i+\frac{1}{2}}^n$  are column vectors of size  $N_r(P^r + 1)$ :

$$\mathbf{Q}_{i+\frac{1}{2}}^{n+1} = \begin{pmatrix} \hat{q}_{i+\frac{1}{2},1}^{n+1} \\ \dots \\ \hat{q}_{i+\frac{1}{2},k}^{n+1} \\ \dots \\ \hat{q}_{i+\frac{1}{2},N_r}^{n+1} \end{pmatrix} \quad \mathbf{G}_{i+\frac{1}{2}}^n = \begin{pmatrix} \hat{G}_{i+\frac{1}{2},1}^n \\ \dots \\ \hat{G}_{i+\frac{1}{2},k}^n \\ \dots \\ \hat{G}_{i+\frac{1}{2},N_r}^n \end{pmatrix} \quad \mathbf{A}_r \mathbf{r}_{i+\frac{1}{2}}^n = \begin{pmatrix} \mathbf{A}_r \mathbf{r}_{i+\frac{1}{2},1}^n \Delta r_{i+\frac{1}{2},1} \\ \dots \\ \mathbf{A}_r \mathbf{r}_{i+\frac{1}{2},k}^n \Delta r_{i+\frac{1}{2},k} \\ \dots \\ \mathbf{A}_r \mathbf{r}_{i+\frac{1}{2},N_r}^n \Delta r_{i+\frac{1}{2},N_r} \end{pmatrix}. \tag{3.55}$$

The unknown vector  $\mathbf{Q}_{i+\frac{1}{2}}^{n+1}$  is isolated and the discrete momentum equations becomes

$$\mathbf{Q}_{i+\frac{1}{2}}^{n+1} = [\mathbf{D}^{-1} \cdot \mathbf{G}]_{i+\frac{1}{2}}^n - \theta \frac{\Delta t}{\Delta x} [\mathbf{D}^{-1} \cdot \mathbf{A}_r]_{i+\frac{1}{2}}^n (p_{i+1}^{n+1} - p_i^{n+1}) \tag{3.56}$$

where  $[\mathbf{D}^{-1} \cdot \mathbf{G}]_{i+\frac{1}{2}}^n = (\mathbf{D}_{i+\frac{1}{2}}^n)^{-1} \cdot \mathbf{G}_{i+\frac{1}{2}}^n$  and  $[\mathbf{D}^{-1} \cdot \mathbf{A}_r]_{i+\frac{1}{2}}^n = (\mathbf{D}_{i+\frac{1}{2}}^n)^{-1} \cdot \mathbf{A}_r \mathbf{r}_{i+\frac{1}{2}}^n$  are computed using the Thomas algorithm extended to block three-diagonal systems. In this way, the exact solution of the system is reached with linear cost in  $N_r$  and it is possible to avoid the use of iterative algorithms like GMRES, which could become inefficient for ill-conditioned problems. Finally, the degrees of freedom for the total mass flow  $Q = \int_0^R 2\pi \rho u r dr$  can be computed by the

scalar product

$$\hat{Q}_{i+\frac{1}{2}}^{n+1} = \underset{\sim}{Aq}_{i+\frac{1}{2}} \cdot \underset{\sim}{Q}_{i+\frac{1}{2}}^{n+1}, \quad (3.57)$$

where  $\underset{\sim}{Aq}_{i+\frac{1}{2}}$  is the following column vector

$$\underset{\sim}{Aq}_{i+\frac{1}{2}} = \begin{pmatrix} \mathbf{Aq}_{i+\frac{1}{2},1} \Delta r_{i+\frac{1}{2},1} \\ \dots \\ \mathbf{Aq}_{i+\frac{1}{2},k} \Delta r_{i+\frac{1}{2},k} \\ \dots \\ \mathbf{Aq}_{i+\frac{1}{2},N_r} \Delta r_{i+\frac{1}{2},N_r} \end{pmatrix} \quad \text{with} \quad \mathbf{Aq}_{i+\frac{1}{2},k} = \int_0^1 \boldsymbol{\varphi}^r(\xi) d\xi. \quad (3.58)$$

### 3.2.5 DG scheme for the $2D_{xr}$ model - Momentum equation - Part B

In the previous section, the momentum equation (3.32b) has been integrated using  $\phi^x = 1$  so that a classical finite difference approach has been adopted for the discretization of the pressure gradient. In the previous situation the cells that belong to the longitudinal mesh were characterized by a single degree of freedom located in the cell barycentre. Now, there are  $P^x + 1$  degrees of freedom in every longitudinal control volume and the integration of Eq. (3.32b) is carried out using an arbitrary polynomial degree for the basis function  $\phi^x$ . In a first stage, the integration is done only in time and in  $r$  direction. This operation yields the following semi-discrete method

$$\mathbf{D}^n \underset{\sim}{Q}^{n+1} = \underset{\sim}{G}^n - \Delta t \underset{\sim}{Ar}^n \frac{\partial p}{\partial x}, \quad (3.59)$$

where we have dropped the spatial subscripts since the axial coordinate  $x$  is still continuous. Now, it is possible to integrate Eq. (3.59) in  $x$  direction

$$\int_{x_i}^{x_{i+1}} \psi^x \left( \mathbf{D}_{i+\frac{1}{2}}^n \underset{\sim}{Q}_{i+\frac{1}{2}}^{n+1} \right) dx = \int_{x_i}^{x_{i+1}} \psi^x \left( \underset{\sim}{G}_{i+\frac{1}{2}}^n - \Delta t \underset{\sim}{Ar}_{i+\frac{1}{2}}^n \frac{\partial p}{\partial x} \right) dx, \quad (3.60)$$

which can be rearranged as

$$\begin{aligned} \hat{Q}_{i+\frac{1}{2}}^{n+1} &= [\widehat{\mathbf{D}^{-1} \mathbf{G}}]_{i+\frac{1}{2}}^n \\ &- \theta \frac{\Delta t}{\Delta x} (\mathbf{M}_{\mathbf{e}}^x)^{-1} \cdot (\mathbf{R}_{\mathbf{p}}^x [\widehat{\mathbf{D}^{-1} \mathbf{A} \mathbf{r}}]_{i+\frac{1}{2}}^n \cdot \hat{\mathbf{p}}_{i+1}^{n+1} - \mathbf{L}_{\mathbf{p}}^x [\widehat{\mathbf{D}^{-1} \mathbf{A} \mathbf{r}}]_{i+\frac{1}{2}}^n \cdot \hat{\mathbf{p}}_i^{n+1}), \end{aligned} \quad (3.61)$$

where the contribution of the explicit terms is collected in the operator  $\hat{\mathbf{G}}_{i+\frac{1}{2}}^n$

$$\hat{\mathbf{G}}_{i+\frac{1}{2}}^n = \widehat{\mathbf{FQ}}_{i+\frac{1}{2}}^n - (1 - \theta) \frac{\Delta t}{\Delta x} (\mathbf{R}_{\mathbf{p}}^x \widehat{\mathbf{A} \mathbf{r}}_{i+\frac{1}{2}}^n \cdot \hat{\mathbf{p}}_{i+1}^n - \mathbf{L}_{\mathbf{p}}^x \widehat{\mathbf{A} \mathbf{r}}_{i+\frac{1}{2}}^n \cdot \hat{\mathbf{p}}_i^n). \quad (3.62)$$



In Eq. (3.61)  $[\widehat{\mathbf{D}^{-1}\mathbf{G}}]_{i+\frac{1}{2}}^n$  and  $[\widehat{\mathbf{D}^{-1}\mathbf{A}\mathbf{r}}]_{i+\frac{1}{2}}^n$  are matrices composed by  $P^x + 1$  column vectors of length  $N_r(P^r + 1)$ , whose components are computed for each of the  $P^x + 1$  axial degrees of freedom using the block Thomas algorithm in radial direction according to the method described in the previous sub section. Then, it is possible to compute the total mass flow using the tensor product in Eq. (3.57) applied for all the  $P^x + 1$  degrees of freedom in the cell  $\Omega_{i+\frac{1}{2}}$ :

$$\begin{aligned} \hat{\mathbf{Q}}_{i+\frac{1}{2}}^n &= [\mathbf{A}\mathbf{q}\widehat{\mathbf{D}^{-1}\mathbf{G}}]_{i+\frac{1}{2}}^n \\ &- \theta \frac{\Delta t}{\Delta x} (\mathbf{M}_e^x)^{-1} \cdot (\mathbf{R}_p^x [\mathbf{A}\mathbf{q}\widehat{\mathbf{D}^{-1}\mathbf{A}\mathbf{r}}]_{i+\frac{1}{2}}^n \cdot \hat{\mathbf{p}}_{i+1}^{n+1} - \mathbf{L}_p^x [\mathbf{A}\mathbf{q}\widehat{\mathbf{D}^{-1}\mathbf{A}\mathbf{r}}]_{i+\frac{1}{2}}^n \cdot \hat{\mathbf{p}}_i^{n+1}). \end{aligned} \quad (3.63)$$

In the last equation  $[\mathbf{A}\mathbf{q}\widehat{\mathbf{D}^{-1}\mathbf{G}}]_{i+\frac{1}{2}}^n$  and  $[\mathbf{A}\mathbf{q}\widehat{\mathbf{D}^{-1}\mathbf{A}\mathbf{r}}]_{i+\frac{1}{2}}^n$  are vectors of size  $P^x + 1$ . For example the  $l$ -th component is given by the scalar product  $[\mathbf{A}\mathbf{q}\widehat{\mathbf{D}^{-1}\mathbf{G}}]_{i+\frac{1}{2},l}^n = \mathbf{A}\mathbf{q}_{i+\frac{1}{2},l}^n \cdot [\widehat{\mathbf{D}^{-1}\mathbf{G}}]_{i+\frac{1}{2},l}^n$ . Casting Eq. (3.63) into the discrete integral continuity equation of Eq. (3.39) gives the following final system for the discrete pressure:

$$\begin{aligned} &\mathbf{M}_e^x \cdot \rho \mathbf{A}(\hat{\mathbf{p}}_i^{n+1}) \\ &- \theta^2 \frac{\Delta t^2}{\Delta x^2} \mathbf{R}_q^x \cdot (\mathbf{M}_e^x)^{-1} \cdot \mathbf{R}_p^x [\mathbf{A}\mathbf{q}\widehat{\mathbf{D}^{-1}\mathbf{A}\mathbf{r}}]_{i+\frac{1}{2}}^n \cdot \hat{\mathbf{p}}_{i+1}^{n+1} \\ &- \theta^2 \frac{\Delta t^2}{\Delta x^2} \mathbf{L}_q^x \cdot (\mathbf{M}_e^x)^{-1} \cdot \mathbf{L}_p^x [\mathbf{A}\mathbf{q}\widehat{\mathbf{D}^{-1}\mathbf{A}\mathbf{r}}]_{i-\frac{1}{2}}^n \cdot \hat{\mathbf{p}}_{i-1}^{n+1} \\ &+ \theta^2 \frac{\Delta t^2}{\Delta x^2} \left( \mathbf{R}_q^x \cdot (\mathbf{M}_e^x)^{-1} \cdot \mathbf{L}_p^x [\mathbf{A}\mathbf{q}\widehat{\mathbf{D}^{-1}\mathbf{A}\mathbf{r}}]_{i+\frac{1}{2}}^n + \mathbf{L}_q^x \cdot (\mathbf{M}_e^x)^{-1} \cdot \mathbf{R}_p^x [\mathbf{A}\mathbf{q}\widehat{\mathbf{D}^{-1}\mathbf{A}\mathbf{r}}]_{i-\frac{1}{2}}^n \right) \cdot \hat{\mathbf{p}}_i^{n+1} = \\ &\mathbf{M}_e^x \cdot \rho \mathbf{A}(\hat{\mathbf{p}}_i^n) - (1 - \theta) \frac{\Delta t}{\Delta x} \left( \mathbf{R}_q^x \cdot \hat{\mathbf{Q}}_{i+\frac{1}{2}}^{n+1} - \mathbf{L}_q^x \cdot \hat{\mathbf{Q}}_{i-\frac{1}{2}}^{n+1} \right) \\ &- \theta \frac{\Delta t}{\Delta x} \left( \mathbf{R}_q^x \cdot (\mathbf{M}_e^x)^{-1} \cdot [\mathbf{A}\mathbf{q}\widehat{\mathbf{D}^{-1}\mathbf{G}}]_{i+\frac{1}{2}}^n - \mathbf{L}_q^x \cdot (\mathbf{M}_e^x)^{-1} \cdot [\mathbf{A}\mathbf{q}\widehat{\mathbf{D}^{-1}\mathbf{G}}]_{i-\frac{1}{2}}^n \right). \end{aligned} \quad (3.64)$$

Eq. yields a mildly nonlinear system of the same form of Eq. (3.25) that can be solved with the same Newton-type method of Casulli et al. Finally, when the pressure is known, the  $P^x + 1$  column vectors  $\hat{\mathbf{Q}}_{i+\frac{1}{2}}^{n+1}$  are computed from Eq. (3.61).

When the longitudinal velocity field  $u$  is known, it is possible to calculate the radial velocity  $w$  from the continuity equation (Eq. (3.32a)). The radial velocity  $w$  is needed for the discretization of the nonlinear convective terms, which is again performed with an explicit RKDG scheme based on the third order TVD Runge-Kutta method, as for the 1D model, but now using the two-dimensional  $x - r$  grid.

When the nonlinear convective terms can be neglected, one can simply set  $\widehat{\mathbf{F}}_{i,k}^n = \hat{\mathbf{q}}_{i,k}^n$  and the scheme becomes unconditionally stable; in this case the time step size  $\Delta t$  has to be chosen properly dependent on the physical problem and small enough to reduce reasonably the numerical viscosity due to the low order time discretization. Moreover, when  $P^x = P^r = 0$  the scheme reduces to the finite volume method investigated in [56, 91] and presented in the previous chapter. In addition, if  $P^r = 0$  and  $N_r = 1$  one gets the high order DG scheme for the 1D model explained previously in section 3.1.

### 3.3 Numerical tests

In the following, the new semi-implicit staggered DG schemes for compressible flows in compliant tubes are validated by comparing the obtained numerical results against different reference solutions available in the literature.

#### 3.3.1 Steady flow in an elastic pipe and numerical convergence study

In the first test considered, see also [34, 56], the flow regime is steady, the fluid is incompressible, the wall displacement is governed by the Laplace law and the convective terms  $u \frac{\partial u}{\partial x} + w \frac{\partial u}{\partial r}$  are neglected. Consequently, the mass flow  $Q$  within the tube is constant and the speed of sound of the fluid tends to infinity so that the density becomes constant and equal to  $\rho_0$ . We stress that this very low Mach number regime close to the incompressible limit is very difficult to simulate for many compressible flow solvers. Under the above assumptions, the analytical solution for the radius in  $x$  direction reads

$$R(x) = \sqrt[5]{R_0^5 - \frac{40\mu Q}{\rho_0 \pi \beta} x}, \quad (3.65)$$

while the axial velocity is computed as

$$u(x, r) = \frac{2Q}{\rho_0 \pi R^4(x)} [R(x)^2 - r^2]. \quad (3.66)$$

Moreover, with the known radius from (3.65), the analytical solution for the pressure is given by

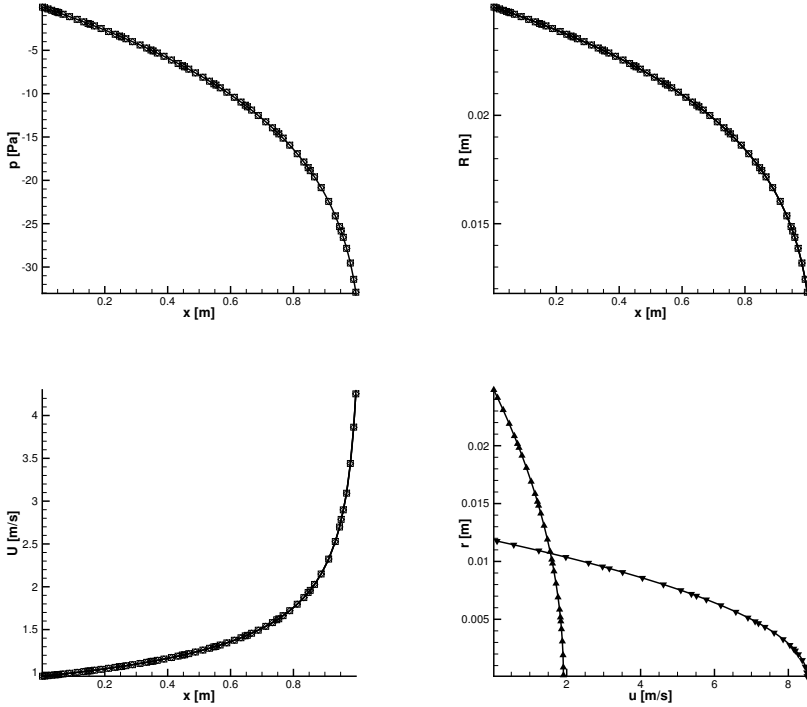
$$p(x) = p_0 - \beta[R(x) - R_0], \quad (3.67)$$

and the mean cross sectional velocity is  $U(x) = Q/[\rho_0 \pi R(x)^2]$ . Numerically, this solution is obtained after a transient period starting from an initial condition at  $t = 0$  when the pressure and the radius are constant,  $p(x, t = 0) = p_0$  and  $R(x, t = 0) = R_0$ , and the velocity fields is  $u(x, r, t = 0) = 0$ . For this test, the geometrical and physical parameters are set as follows:  $L = 1\text{m}$ ,  $R_0 = 0.025\text{ m}$ ,  $Q_0 = 0.001875\text{ kg}^3/\text{s}$ ,  $\rho_0 = 1\text{ kg/m}^3$  and  $\mu = 10^{-3}\text{ Pa}\cdot\text{s}$ . To model a nearly incompressible fluid, the speed of sound is set to  $c_0 = 10^6\text{ m/s}$ , while the stiffness parameter of the tube is set to  $\beta = 2500$ , as in [34]. The simulations are carried out until a final time of  $t_e = 25\text{ s}$  so that a steady state solution has been reached. We use  $N_t = 50$  steps, characterized by a constant time step size equal to  $\Delta t = t_e/N_t = 0.5\text{ s}$  setting  $\theta = 1$ . For the 1D model, the polynomial degree is  $P = 5$  and the cells of the main grid are  $N_x = 10$ . Then, for the  $2D_{x,r}$  model the polynomial degrees of the basis functions are chosen as  $P^x = P^r = 5$ . The main longitudinal grid is composed of  $N_x = 10$  cells while the main radial grid has 7 rings ( $N_r = 7$ ). The results are exposed in Fig. 3.2, which shows an excellent agreement between the numerical data and the exact solution. Moreover, for this smooth problem, for which the analytical solution is known, we also compute the numerical convergence rates shown in tables Tab. 3.1 and Tab. 3.2 with respect to the  $L_2$  norms defined as

$$\epsilon_{L_2}^P = \left( \int_{\Omega} (p_h - p_{ex})^2 dx dr \right)^{\frac{1}{2}} \quad \epsilon_{L_2}^u = \left( \int_{\Omega} (u_h - u_{ex})^2 dx dr \right)^{\frac{1}{2}}. \quad (3.68)$$

For both schemes, i.e. the one applied to the 1D model and to the  $2D_{x,r}$  model, it is possible to state that the order of accuracy for the velocity is  $P$ , while the order of convergence for the pressure

is  $P$  for odd polynomial degrees and  $P + 1$  for even polynomial degrees. This behavior for the pressure convergence has also been observed in the spectral semi-implicit DG schemes developed in [70] for the incompressible Navier-Stokes equations. Further theoretical analysis on this subject will be carried out in the future.



**Figure 3.2:** Steady flow in an elastic pipe. From top left to bottom right: exact solution (—) and numerical data for pressure, radius and mean velocity computed using the the semi-implicit staggered DG scheme for the 1D model ( $\square$ ) and for the 2D<sub>xr</sub> model ( $\diamond$ ). The radial velocity profiles in the last figure can be computed only using the 2D<sub>xr</sub> model and are shown for the inlet of the pipe ( $\blacktriangle$ ) and the outlet of the pipe ( $\blacktriangledown$ ).

$P$	$N_x$	$\epsilon_{L_2}^P$	$\epsilon_{L_2}^U$	$O_{L_2}^P$	$O_{L_2}^U$
1	100	5.8460E-2	1.5546E-2	-	-
1	200	2.9397E-2	7.8068E-3	0.99	0.99
1	400	1.4718E-2	3.9075E-3	0.99	0.99
1	600	9.8148E-3	2.6054E-3	0.99	0.99
2	200	2.4600E-5	1.9169E-4	-	-
2	400	3.0497E-6	3.7170E-5	3.01	2.36
2	600	8.9226E-7	1.3915E-5	3.03	2.42
2	800	3.7302E-7	6.8856E-6	3.03	2.44
3	50	1.5697E-3	3.8997E-4	-	-
3	100	2.0179E-4	5.2540E-5	2.95	2.89
3	200	2.3720E-5	5.9964E-6	3.08	3.13
3	400	2.7655E-6	6.2585E-7	3.10	3.26
4	50	3.9505E-5	8.0151E-5	-	-
4	100	1.8913E-6	6.3574E-6	4.38	3.65
4	200	6.8897E-8	3.9150E-7	4.77	4.02
4	400	2.2076E-9	2.0622E-8	4.96	4.24
5	25	8.8336E-4	1.2298E-4	-	-
5	50	5.0042E-5	8.0968E-6	4.14	3.92
5	100	1.8626E-6	3.5388E-7	4.77	4.51
5	200	5.4435E-8	1.1597E-8	4.96	4.93
6	100	2.4623E-8	4.1609E-8	-	-
6	125	5.8918E-9	1.1804E-8	6.40	5.64
6	150	1.7979E-9	4.1150E-9	6.50	5.78
6	175	7.1047E-10	1.6638E-9	6.02	5.87

**Table 3.1:** Numerical convergence rates computed with respect to the  $L_2$  error norms of pressure and velocity for the semi-implicit staggered DG scheme applied to the 1D model.

### 3.3.2 Womersley profile

For this test case, an exact solution for three-dimensional axially symmetric, incompressible unsteady flows in a rigid pipe is available and was first found by Womersley in [161]. Imposing at

$P^x = P^r$	$N_x = N_r$	$\epsilon_{L_2}^P$	$\epsilon_{L_2}^u$	$O_{L_2}^P$	$O_{L_2}^u$
1	100	5.1619E-2	5.9548E-4		
1	200	2.5974E-2	2.9779E-4	0.99	0.99
1	400	1.7337E-2	1.9840E-4	0.99	1.00
1	600	1.3009E-3	1.4874E-4	0.99	1.00
2	100	1.8383E-4	3.0765E-5		
2	150	5.6960E-5	1.2410E-5	2.88	2.23
2	200	2.4373E-5	6.3846E-6	2.95	2.31
2	250	1.2522E-5	3.7781E-6	2.98	2.35
3	50	4.0088E-4	4.3406E-6		
3	100	1.6740E-4	1.8264E-6	3.03	3.00
3	200	8.4235E-5	9.1593E-7	3.07	3.09
3	400	4.7891E-5	5.1923E-7	3.09	3.14
4	50	3.8172E-5	2.7036E-6		
4	100	6.7817E-6	6.2631E-7	4.26	3.6
4	200	1.8560E-6	2.0892E-7	4.50	3.81
4	400	6.5690E-7	8.6600E-8	4.65	3.94
5	25	4.1763E-5	2.8536E-7		
5	50	6.2875E-6	4.6914E-8	4.66	4.45
5	100	1.5348E-6	1.2154E-8	4.90	4.69
5	200	4.9979E-7	4.1209E-9	5.02	4.84
6	100	1.5287E-6	5.3068E-8		
6	125	1.4321E-7	6.4976E-9	5.83	5.17
6	150	2.4135E-8	1.3418E-9	6.18	5.48
6	175	5.7749E-9	3.7840E-10	6.40	5.67

**Table 3.2:** Numerical convergence rates computed with respect to the  $L_2$  error norms of pressure and velocity for the staggered semi-implicit DG scheme applied to the  $2D_x r$  model.

the inlet of the pipe an oscillating pressure gradient characterized by an amplitude  $A_w$  and by an angular frequency  $\omega = 2\pi f$

$$p_{in}(t) = \Re(p_{out} + LA_w e^{j\omega t}) = p_{out} + LA_w \cos(\omega t), \quad (3.69)$$

with  $\Re$  denoting the real part of the complex number, then the exact solution for the velocity reads

$$u(r, t) = \Re \left( A_w \frac{1}{j\omega} \left[ 1 - \frac{J_0(\hat{r} j^{3/2} Wo)}{J_0(j^{3/2} Wo)} \right] e^{j\omega t} \right). \quad (3.70)$$

In Eq. (3.70)  $J_0$  is the zero-th order complex Bessel function,  $j = \sqrt{-1}$  is the imaginary unit and  $Wo = R\sqrt{\rho w/\mu}$  is the Womersley number. Here we only use the  $2D_{xr}$  model for two simulations characterized by different Womersley regimes. Both numerical experiments share the following numerical and physical parameters:  $\rho_0 = 1000 \text{ kg/m}^3$ ,  $R_0 = 4 \cdot 10^{-3} \text{ m}$ ,  $N_x = 10$  and  $P^5 = 5$ . The stiffness parameter of the pipe is set to  $\beta = 10^{20}$  and the sound speed of the fluid is set to  $c_0 = 10^7$ , as in [56]. Moreover, for this unsteady test, it is crucial to impose  $\theta = 0.52$  in order to improve time accuracy compared to a simple backward Euler scheme and consequently to have a good matching with the reference solution. Furthermore, in both cases the time step  $\Delta t$  is equal to  $T_w/100$  where  $T_w$  is the wave period equal to  $f^{-1}$ . In the first simulation (Test A), the frequency  $f$  is equal to 200 Hz and the viscosity  $\mu$  is  $10^{-1} \text{ Pa}\cdot\text{s}$ . Here, the polynomial degree in radial direction  $P^r$  is 5 and the tube is divided into  $N_r = 7$  rings of constant thickness. From Fig. 3.3 (left) it is possible to see that the agreement between the numerical data and the analytical solution is perfect.

In the second case (Test B), a very high Womersley regime is considered, increasing the wave frequency to  $f = 2000 \text{ Hz}$  and reducing the dynamic viscosity to  $\mu = 10^{-3} \text{ Pa}\cdot\text{s}$ . To make the problem more challenging, we reduce the number of layers to  $N_r = 3$ , but the polynomial degree in radial direction is increased to  $P^r = 10$ . Moreover, since the boundary layer on the wall is very thin, two rings are used to fit the part of the domain close to tube wall bounded between  $0.9R_0$  and  $R_0$ . Fig. 3.3 (right) shows that also in this case the fitting between numerical and analytical solution is very good.

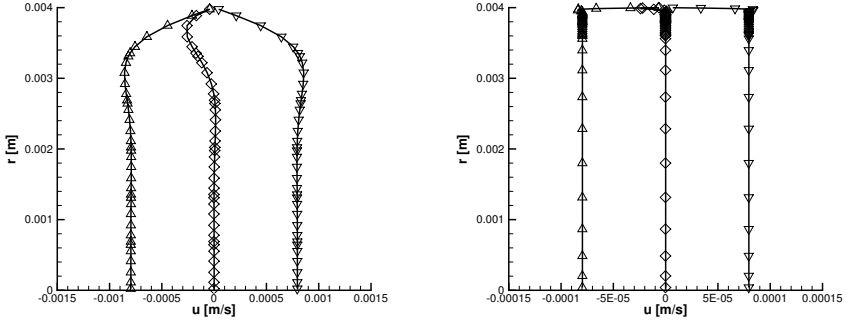
### 3.3.3 Impedance matrix

This test regards the comparison of the numerical solutions against an analytic solution derived in the frequency domain valid for weakly compressible laminar flows and which has first been presented in [105]. The system of equations (3.32) is simplified neglecting the radial velocity  $w$  and the convective terms of the momentum equations. Then one can assume periodic laminar flow in a situation such that the friction effects given by the frequency are higher than those of quasi-steady flow conditions. Using the Laplace transform one obtains the following expression for a hydraulic system composed of a pipe of length  $L$  and radius  $R$

$$\begin{bmatrix} \bar{p}_1(s) \\ \bar{p}_2(s) \end{bmatrix} = \mathbf{Z}(s) \begin{bmatrix} \hat{Q}_1(s) \\ \hat{Q}_2(s) \end{bmatrix} = \begin{bmatrix} Z_L(s) \frac{\cosh(\gamma(s)L)}{\sinh \gamma(s)L} & -Z_L(s) \frac{1}{\sinh(\gamma(s)L)} \\ Z_L(s) \frac{1}{\sinh(\gamma(s)L)} & -Z_L(s) \frac{\cosh(\gamma(s)L)}{\sinh \gamma(s)L} \end{bmatrix} \begin{bmatrix} \bar{Q}_1(s) \\ \bar{Q}_2(s) \end{bmatrix}, \quad (3.71)$$

where  $s = j\omega$  is the Laplace parameter,  $j^2 = -1$ ,  $\bar{p}(x, s)$  and  $\bar{Q}(x, s)$  are the complex pressure and the complex volume flow rate. The indices 1 and 2 indicate the quantities at the inlet and at the outlet of the tube, respectively. Consequently,  $\bar{p}_1 = \bar{p}(0, s)$ ,  $\bar{p}_2 = \bar{p}(L, s)$  and  $\hat{Q}_1 = \hat{Q}(0, s)$  and  $\hat{Q}_2 = \hat{Q}(L, s)$ .  $\mathbf{Z}(s)$  is the impedance matrix and its elements are expressed in terms of the hyperbolic sinus and cosinus functions. Moreover,  $Z_L$  is a transfer coefficient that reads

$$Z_L(s) = Z_0 \sqrt{\frac{J_0(R^*)}{J_1(R^*)}}, \quad \text{with} \quad Z_0 = \frac{\sqrt{E'} \rho}{\pi R^2} \quad \text{and} \quad E' = E \frac{1}{1 + \frac{E}{E_\infty} W}, \quad (3.72)$$



**Figure 3.3:** Womersley profiles for Test A (left) and Test B (right). Analytical solution (—) and numerical data from the  $2D_{xr}$  model after 25 ( $\nabla$ ), 50 ( $\diamond$ ) and 75 ( $\triangle$ ) timesteps for the axial velocity  $u$ .

where  $Z_0$  is the frictionless case of  $Z_L$  and  $E'$  is the modified total bulk modulus of the system that takes into account also the wall elasticity. Here,  $E_\infty$  is the Young modulus of the wall material and  $E$  is the bulk modulus of the fluid, commonly defined as  $E = \rho_0 c_0^2$ ;  $\gamma$  is a wave propagation parameter that is essentially a function of the speed of sound and of the complex radius  $R^*$ :

$$\gamma(s) = \frac{s}{c_0} \sqrt{-\frac{J_0(R^*)}{J_2(R^*)}} \quad \text{with} \quad R^* = j \sqrt{\frac{s}{\nu}} R. \quad (3.73)$$

The elements of the impedance matrix are computed starting from a numerical simulation in the time domain where a frequency is fixed and an oscillatory pressure gradient is imposed, exactly like in the Womersley test. The complex quantities  $\bar{p}$  and  $\bar{Q}$  are given applying the fast Fourier transformation to the time signals of the numerical pressure and mass flow rates taken at the inlet and at the outlet of the tube. Moreover, the impedance matrix is symmetric and consequently the elements  $z_{11}$  and  $z_{12}$  of the impedance matrix can be computed as follows [105]:

$$z_{11}(s) = \frac{\bar{p}_1(s)\bar{Q}_1(s)}{\bar{Q}_1^2(s) - \bar{Q}_2^2(s)}, \quad \text{and} \quad z_{12}(s) = -\frac{\bar{p}_1(s)\bar{Q}_2(s)}{\bar{Q}_1^2(s) - \bar{Q}_2^2(s)}. \quad (3.74)$$

At first, the impedance matrix for a test fluid has been computed for a pipe assumed to be rigid. Here, the length of the tube is  $L = 1$  m, the radius  $R_0 = 4$  mm. The reference density is  $\rho_0 = 998.2$  kg/m<sup>3</sup>, the speed of sound is equal to  $c_0 = 1400$  m/s and the dynamic viscosity is  $\mu = 10^{-2}$  Pa·s. Then, the elastic case is investigated. The physical and geometrical parameters are the same considered for the previous case except for the viscosity  $\mu$  that now is equal to  $10^{-1}$  Pa·s and the elasticity parameters. We set  $E_\infty = 10^{10}$  Pa and the wall parameter  $W$  in Eq. (3.4)

is computed as follows

$$W = 2 \frac{\left(\frac{h_w}{R_0}\right)^2 (1 + \nu_w) - (1 - 2\nu_w)}{\left(\frac{h_w}{R_0}\right)^2 - 1}, \quad (3.75)$$

where  $h_w = 1$  mm is the wall thickness and  $\nu_w = 0.5$  is the Poisson ratio of the wall material. For more details about the physical meaning of the impedance matrix see [91, 105]. In both simulations, the DG scheme for the 1D model has been applied using 100 cells ( $N_x = 100$ ), with a polynomial degree  $P = 5$ , while for the DG scheme applied to the  $2D_{xr}$  model the number of elements are  $N_x = 100$  and  $N_r = 20$  and the polynomial degrees used are  $P^x = 5$  and  $P^r = 4$ . The results for the elements  $z_{11}$  and  $z_{12}$  of the impedance matrix for rigid and elastic pipe are exposed in Fig. 3.4 and in Fig. 3.5, respectively. In all the plots the agreement between numerical and analytical solution is very good.

### 3.3.4 Blasius boundary layer

Here, we use the semi-implicit staggered DG scheme in order to investigate the boundary layer effects for laminar high Reynolds number flows. The Blasius boundary layer solution [12, 121] is defined by the ODE

$$f''' + ff' = 0, \quad (3.76)$$

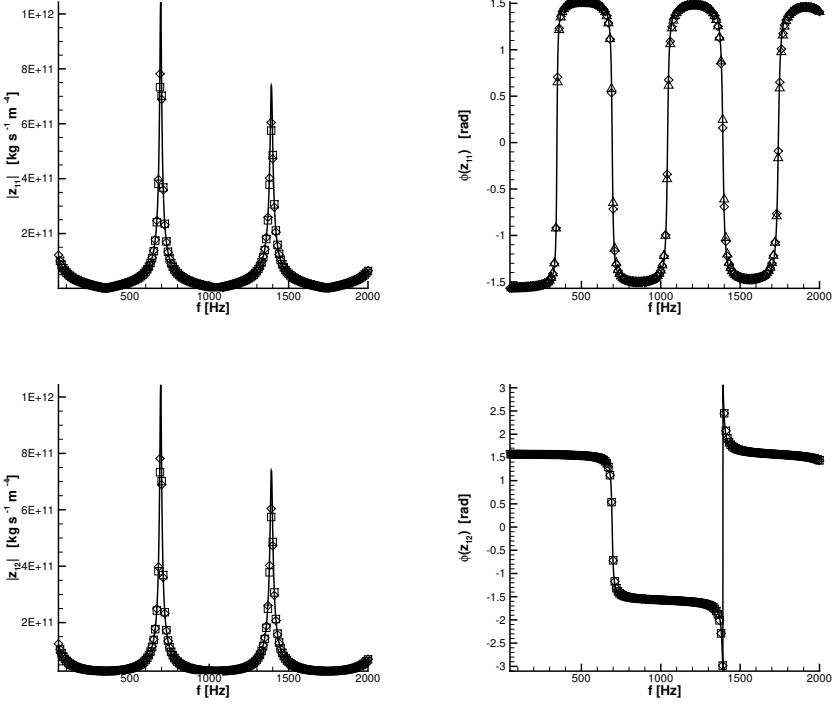
with the boundary conditions

$$f(0) = 0, \quad f'(0) = 0, \quad \lim_{\eta \rightarrow \infty} f'(\eta) = 1 \quad \text{with} \quad \eta = (R_0 - r) \sqrt{\frac{U_\infty}{2\nu x}}, \quad (3.77)$$

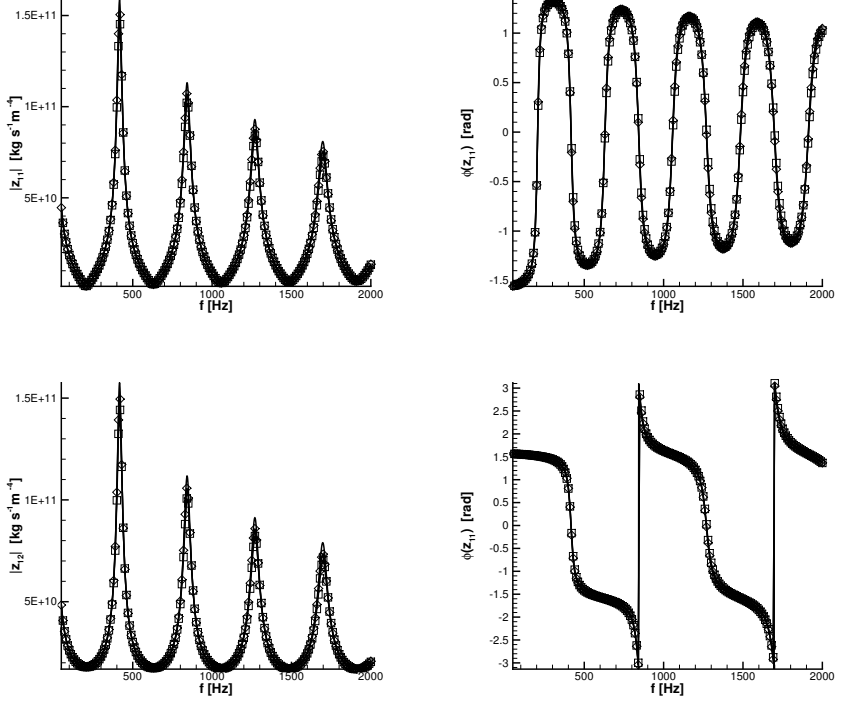
where  $\eta$  is the Blasius similarity coordinate and  $U_\infty$  is the undisturbed velocity (see [128]). The solution of Eq. (3.76) is found using the shooting method and the velocity profile is equal to  $u(x, y) = U_\infty f'(\eta)$ .

In this test, the role of the non-linear convective terms is crucial. The test is carried out using a straight rigid tube of radius  $R_0 = 0.1$  m and of length  $L = 1$  m. The initial condition for the pressure is  $p(x, 0) = p_0 = 0$  Pa and for the velocity is  $u(x, r, 0) = U_0 = 1$  m/s. The fluid is assumed to be almost incompressible ( $c_0 = 10^6$ ) and consequently the density tends to be constant and equal to  $\rho_0 = 1$  kg/m<sup>3</sup>. The dynamic viscosity is set to  $\mu = 10^{-5}$  Pa·s, although in this Reynolds number regime the flow would not be laminar any more. The domain is discretized with 250 cells in longitudinal direction and by 100 rings in the radial direction ( $N_x = 250$  and  $N_r = 100$ ). At the inlet of the pipe the velocity is imposed equal to  $u(0, t) = U_0$  while at the outlet the pressure is  $p(L, t) = p_0$ . The parameter  $\theta$  can be set equal to 1, since we are interested in obtaining a steady solution. The polynomial degree of the basis functions in  $r$  direction is  $P^r = 2$ . We decided to use  $P^x = 0$ , since the most important physical effects are due to the derivatives in radial direction. Numerically, a transient phase occurs, then for this set of parameters a steady state regime is achieved at  $t = 2.5$  s. The radial velocity profiles are taken at  $x = 0.75$  m. Moreover, since it is known that with a constant pipe radius and a rigid wall the inner core of the fluid accelerates (due to mass conservation and the so-called displacement thickness of the boundary layer), the comparison against the reference solution is done by setting the velocity  $U_\infty$  to the numerical velocity obtained in the fluid core at  $r = 0$ . The plot in Fig. 3.6 shows an excellent fitting between numerical data and the Blasius solution. Consequently, we can state that in our approach also the radial velocity  $w$  and the non-linear convective terms are computed properly.

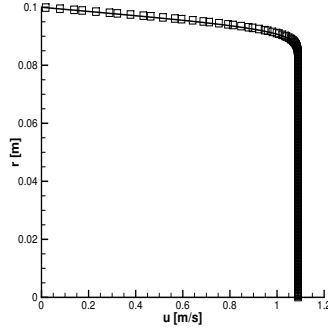




**Figure 3.4:** Absolute value and phase of the elements  $z_{11}$  and  $z_{12}$  of the impedance matrix for a fluid in a rigid pipe - Reference solution (—) and numerical data using the semi-implicit staggered DG scheme applied to the 1D model ( $\square$ ) and to the  $2D_{xr}$  model ( $\diamond$ ).



**Figure 3.5:** Absolute value and phase of the elements  $z_{11}$  and  $z_{12}$  of the impedance matrix for a fluid in an elastic pipe - Reference solution (—) and numerical data using the semi-implicit staggered DG scheme applied to the 1D model ( $\square$ ) and to the  $2D_{x_T}$  model ( $\diamond$ ).



**Figure 3.6:** Blasius boundary layer. Reference solution (—) and numerical results using the staggered SIDG scheme applied to the  $2D_{xr}$  model ( $\square$ ).

### 3.4 Conclusions on staggered semi-implicit DG schemes for pipe flow

We have developed and implemented a new semi-implicit staggered Discontinuous Galerkin scheme for two mathematical models that describe weakly compressible flows in elastic tubes. In particular, we have discretized a cross-sectionally averaged 1D model and a  $2D_{xr}$  model. Both can be derived from the three-dimensional compressible Navier-Stokes equations assuming axial symmetry and hydrostatic flow, i.e. constant pressure within each cross section. The proposed semi-implicit DG finite element approach allows to reach arbitrary high order of accuracy in space, while up to second order of accuracy is achieved in time using the  $\theta$ -method. Both schemes are characterized by a mild CFL condition and consequently the time step  $\Delta t$  depends only on the fluid velocity and not on the sound speed. In addition, when the non-linear convective terms can be neglected, the numerical methods are unconditionally stable. The pressure is efficiently computed solving a mildly non-linear system which is observed to be symmetric and in general well-conditioned. Then, the velocity field can be readily updated by using the discrete momentum equations.

The novel family of numerical methods has been compared against several analytical solutions for steady and unsteady test problems. We have observed an excellent agreement in all the benchmark problems that have been simulated.



## 4 A posteriori sub-cell finite volume limiting of staggered SIDG schemes for the shallow water equations

The high order staggered semi-implicit DG schemes introduced at the end of the previous chapter were still without any sort of nonlinear limiter. And according to Godunov's theorem, any unlimited high order scheme inevitably produces spurious oscillations in the vicinity of discontinuities and strong gradients. In this chapter we therefore propose a novel family of *a posteriori* sub-cell finite volume limiters for spatially high order accurate semi-implicit discontinuous Galerkin (DG) schemes on staggered Cartesian grids for the solution of the shallow water equations expressed in conservative form in one and two space dimensions. To the best of our knowledge, this is the first time that an *a posteriori* subcell FV limiting strategy is applied to *semi-implicit* DG schemes on staggered grids. All previous subcell FV limiters for DG methods were limited to the explicit case so far.

We first start with a description of the unlimited arbitrary high order accurate staggered semi-implicit DG scheme. In this method, the continuity equation and the momentum equations are integrated using a discontinuous finite element strategy on staggered control volumes, where the discrete free surface elevation is defined on the main grid and the discrete momentum is defined on edge-based staggered dual control volumes. As usual, in the semi-implicit approach pressure terms are discretized implicitly, while the nonlinear convective terms are discretized explicitly. Inserting the momentum equations into the discrete continuity equation leads to a well conditioned block penta-diagonal linear system for the free surface elevation which can be efficiently solved with modern iterative methods.

In this chapter, we therefore propose to extend the successful family of *a posteriori* subcell finite volume limiters recently introduced by Dumbser et al. (2014) for explicit DG schemes also to semi-implicit time discretizations. At time  $t^n$  the unlimited DG scheme is run in order to produce a so-called *candidate solution* for time  $t^{n+1}$ . Then, the cells characterized by a non-admissible candidate solution are found by using physical and numerical detection criteria based on the positivity of the solution, the absence of floating point errors and the use of a relaxed discrete maximum principle (DMP) according to the MOOD strategy of Clain, Loubère and Diot (2013). In all the cells that are flagged as troubled control volumes a more robust semi-implicit finite volume (FV) method is then applied on a sub-grid composed of  $2P + 1$  cells, where  $P$  denotes the polynomial degree used for approximating the discrete solution within the DG scheme. Then, after having identified the troubled cells, the linear system for the new free surface elevation is assembled and solved again, where unlimited cells use the high order semi-implicit DG scheme and limited cells are evolved via the more robust finite volume method. Finally, from the subcell finite volume averages a higher order DG polynomial is reconstructed and then the scheme proceeds with the next time step.

We apply the new semi-implicit staggered DG method with a *a posteriori* subcell FV limiter to classical benchmarks such as Riemann problems in 1D and circular dam-break problems in 2D with

shock waves, showing that the new subcell finite volume limiter is able to resolve shocks accurately without producing spurious oscillations. Moreover, if the solution is smooth, the detector does not find any troubled cells, as expected; consequently, the limiter is not activated and the method corresponds to the unlimited staggered semi-implicit DG scheme. In addition, we carry out numerical tests which show that the new scheme is well-balanced and able to deal with wet and dry fronts.

## 4.1 Sub-cell finite volume limiting of staggered semi-implicit DG schemes in 1D

In this section we present our new numerical scheme in the one-dimensional setting, applied to the well-known shallow water equations. We first recall the governing equations and the unlimited high order semi-implicit DG scheme [54] as well as the semi-implicit finite volume method [27, 99, 134], which is then used as subcell FV limiter of the high order DG scheme. Subsequently we describe the *a posteriori* limiter, focusing on the detection criteria and the MOOD algorithm. Moreover, several matrices and tensors will be introduced during the derivation of the schemes, which are explicitly given for some polynomial degrees  $P$  in the appendix A.1.

### 4.1.1 Governing equations of the 1D model

The frictionless shallow water equations form a hyperbolic system that consists of the continuity equation and the momentum equation in  $x$  direction. They read

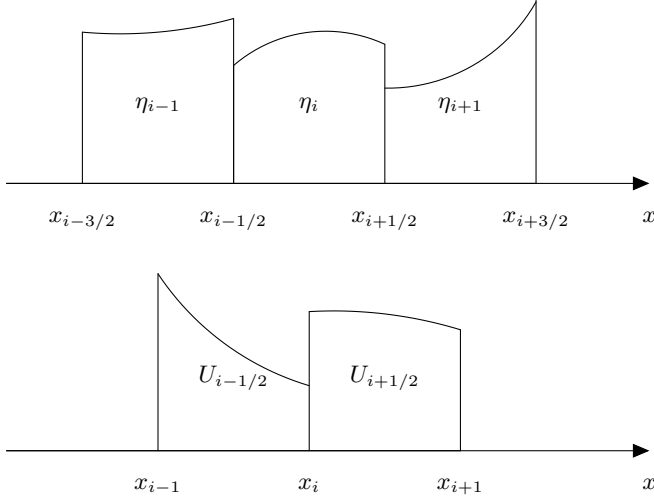
$$\begin{aligned} \frac{\partial \eta}{\partial t} + \frac{\partial U}{\partial x} &= 0, \\ \frac{\partial U}{\partial t} + \frac{\partial uU}{\partial x} + gH \frac{\partial \eta}{\partial x} &= 0, \end{aligned} \tag{4.1}$$

where  $x$  and  $t$  denote the spatial and the temporal coordinates,  $\eta(x, y)$  is the position of free surface,  $h(x)$  is the bottom topography which is considered fixed in time,  $H(x, t) = h(x) + \eta(x, t)$  is the water depth,  $u(x, t)$  is the velocity field,  $U(x, t) = u(x, t)H(x, t)$  is the volumetric discharge and  $g$  is the gravity constant. The computational domain is denoted by  $\Omega_x = [x_L, x_R]$ .

### 4.1.2 Unlimited staggered semi-implicit DG scheme for the 1D shallow water equations

The computational domain  $\Omega_x$  is composed of two overlapping grids (see Fig. 4.1). The first is the so-called main grid which contains  $N_x$  equally spaced elements of length  $\Delta x = \frac{L}{N_x} = \frac{x_R - x_L}{N_x}$ ; a cell that belongs to this grid is denoted by  $T_i = [x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}]$ . The dual mesh has  $N_x + 1$  elements; the edges of the generic control volume  $T_{i+\frac{1}{2}} = [x_i, x_{i+1}]$  are the barycenters of two consecutive cells of the main grid. Hence, on this grid there are  $N_x - 1$  equally spaced elements of length  $\Delta x$ , and the length of the two cells at the left and at the right boundary is  $\Delta x/2$ .

According to the finite element approach, the numerical solution is projected onto piecewise polynomials of degree  $P$ . For this purpose, we introduce a set of basis functions  $\varphi(\xi)$  defined on the



**Figure 4.1:** Staggered grids for the one-dimensional DG scheme. Main grid used for the free surface (top) and dual mesh for the velocity (bottom).

reference element  $[0, 1]$ . Here we use a nodal basis given by the Lagrange interpolation polynomials passing through the Gauss-Legendre quadrature points. In the physical control volumes the basis functions  $\phi_l(x)$  on the main grid and the basis functions  $\psi_l(x)$  on the dual grid are generated from  $\varphi_l(\xi)$  as follows

$$\begin{aligned} \phi_l(x) &= \varphi_l(\xi) \quad \text{with} \quad x = x_i + \xi \Delta x, \quad \text{and} \\ \psi_l(x) &= \varphi_l(\xi) \quad \text{with} \quad x = x_{i+\frac{1}{2}} + \xi \Delta x, \quad 0 \leq \xi \leq 1. \end{aligned} \quad (4.2)$$

The free surface elevation is a quantity located on the main grid

$$\eta_i(x, t) = \sum_{l=1}^{P+1} \phi_l(x) \hat{\eta}_{i,l}(t) := \boldsymbol{\phi}(x) \cdot \hat{\boldsymbol{\eta}}_i(t), \quad (4.3)$$

while  $U$ ,  $u$  and  $H$  are defined on the staggered dual mesh:

$$\begin{aligned} U_{i+\frac{1}{2}}(x, t) &= \sum_{l=1}^{P+1} \psi_l(x) \hat{U}_{i+\frac{1}{2},l}(t) := \boldsymbol{\psi}(\mathbf{x}) \cdot \hat{\mathbf{U}}_{i+\frac{1}{2}}(t), \\ u_{i+\frac{1}{2}}(x, t) &= \boldsymbol{\psi}(\mathbf{x}) \cdot \hat{\mathbf{u}}_{i+\frac{1}{2}}(t), \\ H_{i+\frac{1}{2}}(x, t) &= \boldsymbol{\psi}(\mathbf{x}) \cdot \hat{\mathbf{H}}_{i+\frac{1}{2}}(t). \end{aligned} \quad (4.4)$$

The continuity equation is multiplied by the test functions  $\phi$  and is integrated in space over the control volume of the main grid  $T_i$ :

$$\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \phi \left( \frac{\partial \eta}{\partial t} + \frac{\partial U}{\partial x} \right) dx = 0. \quad (4.5)$$

Integration by parts then yields

$$\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \phi \frac{\partial \eta}{\partial t} dx + \phi(x_{i+\frac{1}{2}})U(x_{i+\frac{1}{2}}) - \phi(x_{i-\frac{1}{2}})U(x_{i-\frac{1}{2}}) - \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \frac{\partial \phi}{\partial x} U dx = 0. \quad (4.6)$$

The spatial derivative is discretized in time by using the  $\theta$ -method where  $\theta$  is an implicitness parameter taken in the range  $0.5 \leq \theta \leq 1$ . This time integration corresponds to the implicit Euler method when  $\theta = 1$  and to a second order Crank-Nicolson type method when  $\theta = 0.5$  [32, 109]. In this last case the method is not monotone for arbitrary time step size  $\Delta t$  and may generate oscillations near discontinuities. In addition, the monotonicity violation for  $\theta = 0.5$  can be justified by Godunov's theorem [78], which states that linear monotone schemes can be at most of order one. For this reason the  $\theta$ -method is formulated so that the implicitness parameter  $\theta$  can assume different values in each control volume. More details for the choice of the parameter  $\theta$  are given in the next subsections. Consequently, we obtain the following compact matrix-vector notation for the discrete continuity equation:

$$\mathbf{M} \cdot (\hat{\boldsymbol{\eta}}_i^{n+1} - \hat{\boldsymbol{\eta}}_i^n) + \frac{\Delta t}{\Delta x} (\mathbf{R}_J^{\text{DG}} \cdot \hat{\mathbf{U}}_{i+\frac{1}{2}}^{n+\theta} - \mathbf{L}_J^{\text{DG}} \cdot \hat{\mathbf{U}}_{i-\frac{1}{2}}^{n+\theta}) = 0. \quad (4.7)$$

where  $\hat{\mathbf{U}}_{i+\frac{1}{2}}^{n+\theta} = (1 - \theta_{i+\frac{1}{2}}) \hat{\mathbf{U}}_{i+\frac{1}{2}}^n + \theta_{i+\frac{1}{2}} \hat{\mathbf{U}}_{i+\frac{1}{2}}^{n+1}$ . Similarly, we multiply the momentum equation by the test functions  $\boldsymbol{\psi}$  and integrate over the control volume  $T_{i+\frac{1}{2}}$  of the dual mesh

$$\int_{x_i}^{x_{i+1}} \boldsymbol{\psi} \left( \frac{\partial U}{\partial t} + \frac{\partial u U}{\partial x} + g H \frac{\partial \eta}{\partial x} \right) dx = 0. \quad (4.8)$$

The resulting discrete momentum equation reads in compact matrix-vector notation (see [54])

$$\mathbf{M} \cdot (\hat{\mathbf{U}}_{i+\frac{1}{2}}^{n+1} - \widehat{\mathbf{F}} \hat{\mathbf{U}}_{i+\frac{1}{2}}^n) + g \frac{\Delta t}{\Delta x} (\mathbf{R}_u^{\text{DG}} \cdot \hat{\mathbf{H}}_{i+\frac{1}{2}}^n \hat{\boldsymbol{\eta}}_{i+1}^{n+\theta} - \mathbf{L}_u^{\text{DG}} \cdot \hat{\mathbf{H}}_{i+\frac{1}{2}}^n \hat{\boldsymbol{\eta}}_i^{n+\theta}) = 0 \quad (4.9)$$



where  $\widehat{\mathbf{F}\mathbf{U}}_{i+\frac{1}{2}}^n$  is an explicit discretization of the nonlinear convective term that will be explained later and where  $\hat{\boldsymbol{\eta}}_i^{n+\theta_{i+\frac{1}{2}}} = (1 - \theta_{i+\frac{1}{2}})\hat{\boldsymbol{\eta}}_i^n + \theta_{i+\frac{1}{2}}\hat{\boldsymbol{\eta}}_i^{n+1}$ . The choice of the parameter  $\theta_{i+\frac{1}{2}}$  for time integration in eq. (4.9) will be clarified in the subsection 4.1.3. We cast all the explicit terms into a new quantity  $\hat{\mathbf{G}}_{i+\frac{1}{2}}^n$  computed as follows

$$\hat{\mathbf{G}}_{i+\frac{1}{2}}^n = \widehat{\mathbf{F}\mathbf{U}}_{i+\frac{1}{2}}^n - g \frac{\Delta t}{\Delta x} (1 - \theta_{i+\frac{1}{2}}) \mathbf{M}^{-1} \cdot (\mathbf{R}_{\mathbf{u}}^{\text{DG}} \cdot \hat{\mathbf{H}}_{i+\frac{1}{2}}^n \hat{\boldsymbol{\eta}}_{i+1}^n - \mathbf{L}_{\mathbf{u}}^{\text{DG}} \cdot \hat{\mathbf{H}}_{i+\frac{1}{2}}^n \hat{\boldsymbol{\eta}}_i^n), \quad (4.10)$$

so that the final form of the momentum equation reads

$$\hat{\mathbf{U}}_{i+\frac{1}{2}}^{n+1} = \hat{\mathbf{G}}_{i+\frac{1}{2}}^n - g \frac{\Delta t}{\Delta x} \theta_{i+\frac{1}{2}} \mathbf{M}^{-1} \cdot (\mathbf{R}_{\mathbf{u}}^{\text{DG}} \cdot \hat{\mathbf{H}}_{i+\frac{1}{2}}^n \hat{\boldsymbol{\eta}}_{i+1}^{n+1} - \mathbf{L}_{\mathbf{u}}^{\text{DG}} \cdot \hat{\mathbf{H}}_{i+\frac{1}{2}}^n \hat{\boldsymbol{\eta}}_i^{n+1}). \quad (4.11)$$

Note that the numerical discretization of eq. (4.11) formally never achieves second order of accuracy in time since the water depth  $\hat{\mathbf{H}}_{i+\frac{1}{2}}$  is taken at time level  $t^n$ . However, in general this approximation gives satisfactory results for simulations in practical applications; for more details see [32]. Substituting the expression of  $\hat{\mathbf{U}}_{i+\frac{1}{2}}^{n+1}$  given by eq. (4.11) into eq. (4.7) yields the following linear block three-diagonal system

$$\begin{aligned} \mathbf{M} \cdot \hat{\boldsymbol{\eta}}_i^{n+1} + g \frac{\Delta t^2}{\Delta x^2} \left[ \theta_{i+\frac{1}{2}}^2 \mathbf{R}_{\mathbf{J}}^{\text{DG}} \mathbf{M}^{-1} \left( \mathbf{R}_{\mathbf{u}}^{\text{DG}} \cdot \hat{\mathbf{H}}_{i+\frac{1}{2}}^n \hat{\boldsymbol{\eta}}_{i+1}^{n+1} - \mathbf{L}_{\mathbf{u}}^{\text{DG}} \cdot \hat{\mathbf{H}}_{i+\frac{1}{2}}^n \hat{\boldsymbol{\eta}}_i^{n+1} \right) \right. \\ \left. - \theta_{i-\frac{1}{2}}^2 \mathbf{L}_{\mathbf{J}}^{\text{DG}} \mathbf{M}^{-1} \left( \mathbf{R}_{\mathbf{u}}^{\text{DG}} \cdot \hat{\mathbf{H}}_{i-\frac{1}{2}}^n \hat{\boldsymbol{\eta}}_i^{n+1} - \mathbf{L}_{\mathbf{u}}^{\text{DG}} \cdot \hat{\mathbf{H}}_{i-\frac{1}{2}}^n \hat{\boldsymbol{\eta}}_{i-1}^{n+1} \right) \right] = \hat{\mathbf{b}}_i^n, \end{aligned} \quad (4.12)$$

or more compactly

$$\mathcal{L}_i^{\text{DG}} \cdot \hat{\boldsymbol{\eta}}_{i-1}^{n+1} + \mathcal{C}_i^{\text{DG}} \cdot \hat{\boldsymbol{\eta}}_i^{n+1} + \mathcal{R}_i^{\text{DG}} \cdot \hat{\boldsymbol{\eta}}_{i+1}^{n+1} = \hat{\mathbf{b}}_i^n, \quad (4.13)$$

where  $\mathcal{L}_i^{\text{DG}}$ ,  $\mathcal{C}_i^{\text{DG}}$ ,  $\mathcal{R}_i^{\text{DG}}$  are the lower diagonal, central and upper diagonal blocks and  $\hat{\mathbf{b}}_i^n$  is the known term at the right hand side computed as follows

$$\begin{aligned} \hat{\mathbf{b}}_i^n = \mathbf{M} \cdot \hat{\boldsymbol{\eta}}_i^n - \frac{\Delta t}{\Delta x} (\theta_{i+\frac{1}{2}} \mathbf{R}_{\mathbf{J}}^{\text{DG}} \cdot \hat{\mathbf{G}}_{i+\frac{1}{2}}^n - \theta_{i-\frac{1}{2}} \mathbf{L}_{\mathbf{J}}^{\text{DG}} \cdot \hat{\mathbf{G}}_{i-\frac{1}{2}}^n) \\ - \frac{\Delta t}{\Delta x} ((1 - \theta_{i+\frac{1}{2}}) \mathbf{R}_{\mathbf{J}}^{\text{DG}} \cdot \hat{\mathbf{U}}_{i+\frac{1}{2}}^n - (1 - \theta_{i-\frac{1}{2}}) \mathbf{L}_{\mathbf{J}}^{\text{DG}} \cdot \hat{\mathbf{U}}_{i-\frac{1}{2}}^n). \end{aligned} \quad (4.14)$$

The linear system (5.28) can be solved by the Thomas algorithm properly adapted for block three-diagonal systems. Once the new free surface elevation is known, it is possible to compute the degrees of freedom of  $\hat{\mathbf{U}}_{i+\frac{1}{2}}^{n+1}$  using Eq. (4.11) and of the water depth  $\hat{\mathbf{H}}_{i+\frac{1}{2}}$ . However, since the degrees of freedom of  $H$  are defined on the staggered dual grid, while the free surface elevation  $\eta$  is defined on the main grid, the following interpolation based on  $L_2$  projection is used:

$$\hat{\mathbf{H}}_{i+\frac{1}{2}} = \hat{h}_{i+\frac{1}{2}} + \mathbf{M}^{-1} \cdot (\mathbf{M}_{\mathbf{L}}^{\text{DG}} \cdot \hat{\boldsymbol{\eta}}_i + \mathbf{M}_{\mathbf{R}}^{\text{DG}} \cdot \hat{\boldsymbol{\eta}}_{i+1}). \quad (4.15)$$

Now the procedure for the computation of the operator  $\mathbf{F}\mathbf{U}$  is explained. It is easier to discretize this term on the main grid, hence the velocity and the volumetric discharge have to be projected on

the main grid using the same approach used in eq. (4.15), i.e.

$$\tilde{U}_i = \mathbf{M}^{-1} \cdot (\mathbf{M}_L^{\text{DG}} \cdot \tilde{U}_{i-\frac{1}{2}} + \mathbf{M}_R^{\text{DG}} \cdot \tilde{U}_{i+\frac{1}{2}}). \quad (4.16)$$

To obtain a stable discretization, a classical explicit TVD Runge Kutta discontinuous Galerkin scheme (RKDG) on the collocated grid is applied, i.e.

$$\begin{aligned} \hat{\mathbf{k}}_i^1 &= \tilde{U}_i^n + L_h(\hat{\mathbf{u}}^n, \tilde{U}_i^n), \\ \hat{\mathbf{k}}_i^2 &= \frac{3}{4}\tilde{U}_i^n + \frac{1}{4}\hat{\mathbf{k}}_i^1 + \frac{1}{4}L_h(\hat{\mathbf{u}}^n, \hat{\mathbf{k}}_i^1), \\ \widehat{\mathbf{F}\mathbf{U}}_i^n &= \frac{1}{3}\tilde{U}_i^n + \frac{2}{3}\hat{\mathbf{k}}_i^2 + \frac{2}{3}L_h(\hat{\mathbf{u}}^n, \hat{\mathbf{k}}_i^2), \end{aligned} \quad (4.17)$$

where the operator  $L_h(\mathbf{u}, U)$  for cell  $T_i$  reads

$$L_h(\mathbf{u}, U) = -\frac{\Delta t}{\Delta x} \mathbf{M}^{-1} \left( \varphi(1)f_{i+\frac{1}{2}} - \varphi(0)f_{i-\frac{1}{2}} - \mathbf{K} \cdot \hat{\mathbf{f}}_i \right). \quad (4.18)$$

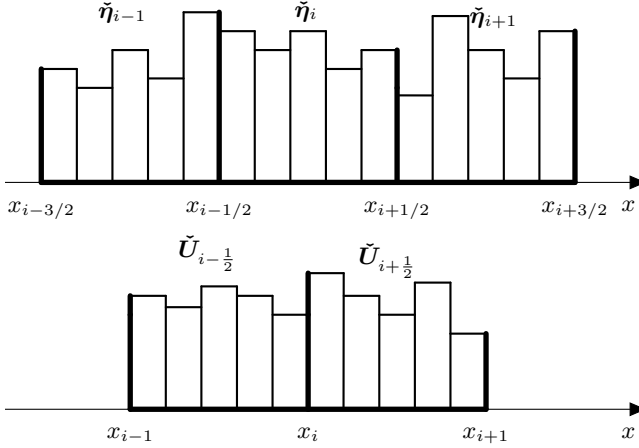
Here,  $\hat{\mathbf{f}}_i = \hat{\mathbf{u}}_i \tilde{U}_i$  and  $f_{i+\frac{1}{2}}$  is a numerical flux at the cell interface given by the Rusanov method as

$$f_{i+\frac{1}{2}} = \frac{1}{2} \left( u_{i+\frac{1}{2}}^+ U_{i+\frac{1}{2}}^+ + u_{i+\frac{1}{2}}^- U_{i+\frac{1}{2}}^- \right) - \frac{1}{2} \max \left( 2|u_{i+\frac{1}{2}}^+|, 2|u_{i+\frac{1}{2}}^-| \right) \left( U_{i+\frac{1}{2}}^+ - U_{i+\frac{1}{2}}^- \right), \quad (4.19)$$

where  $u_{i+\frac{1}{2}}^\pm$  and  $U_{i+\frac{1}{2}}^\pm$  denote the boundary-extrapolated values on the left and right of the interface  $x_{i+\frac{1}{2}}$ , respectively. Finally, the degrees of freedom  $\widehat{\mathbf{F}\mathbf{U}}_{i+\frac{1}{2}}^n$  are projected back to the dual mesh with the same procedure used in eq. (4.15). The maximum admissible time step  $\Delta t$  is given by the usual CFL condition for RKDG schemes, which in this case reads  $\Delta t < \text{CFL} \frac{\Delta x}{2 \max(|u|)}$  with  $\text{CFL} < (2P+1)^{-1}$ . Note that this time step restriction is based only on the fluid velocity and *not* on the surface wave speed  $\sqrt{gH}$ . Moreover, the scheme becomes unconditionally stable if one chooses an explicit Eulerian-Lagrangian DG scheme for the convective terms or if the nonlinear convective terms can be neglected, hence simply setting  $\widehat{\mathbf{F}\mathbf{U}}_{i+\frac{1}{2}}^n := \tilde{U}_{i+\frac{1}{2}}^n$ . For more details about the mathematical properties of high-order staggered semi-implicit schemes see [70, 138, 139, 140] while for the stability analysis of Runge-Kutta TVD methods see [13, 82, 109].

### 4.1.3 A sub-cell formulation for the semi-implicit finite volume method for the 1D shallow water equations

In the following we consider the sub-cell formulation of the semi-implicit finite volume method [27, 99, 134]. Also here the computational domain  $\Omega_x$  is composed of two overlapping meshes according to the staggered philosophy. Considering the element  $T_i$  introduced in the previous section, for a positive integer  $P$ , we use  $2P+1$  piecewise constant subcell averages in order to represent the FV data on  $2P+1$  sub-cells  $T_{i,s}$  of length  $\Delta x_s = \Delta x/(2P+1)$  (see Fig.4.2). As a consequence we have that  $T_i = \cup_s T_{i,s}$  with  $s = 1, \dots, (2P+1)$ . For example, in the case  $P = 2$  the sub-cell degrees of freedom in the cell  $T_i$  are denoted by  $\tilde{\eta}_i = [\tilde{\eta}_{i,1}, \tilde{\eta}_{i,2}, \tilde{\eta}_{i,3}, \tilde{\eta}_{i,4}, \tilde{\eta}_{i,5}]$ .



**Figure 4.2:** Staggered grids for the sub-cell finite volume scheme in one dimension in the case  $P = 2$  ( $N_s = 5$ ). Main grid used for the free surface (top) and staggered dual mesh for the velocity (bottom).

Also for this case the free surface  $\eta$  is located on the main grid and the quantities  $U$ ,  $u$ ,  $H$  are considered on the control volumes of the dual grid. The semi-implicit discrete continuity equation with the adaptive  $\theta$ -method reads

$$\check{\eta}_i^{n+1} - \check{\eta}_i^n + \frac{\Delta t}{\Delta x} (\mathbf{R}^{\text{FV}} \cdot \check{U}_{i+\frac{1}{2}}^{n+\theta_{i+\frac{1}{2}}} - \mathbf{L}^{\text{FV}} \cdot \check{U}_{i-\frac{1}{2}}^{n+\theta_{i-\frac{1}{2}}}) = 0, \quad (4.20)$$

while the discrete momentum equation takes the form

$$(\check{U}_{i+\frac{1}{2}}^{n+1} - \check{F} \check{U}_{i+\frac{1}{2}}^n) + g \frac{\Delta t}{\Delta x} \check{H}_{i+\frac{1}{2}}^n \cdot (\mathbf{R}^{\text{FV}} \cdot \check{\eta}_{i+1}^{n+\theta_{i+\frac{1}{2}}} - \mathbf{L}^{\text{FV}} \cdot \check{\eta}_i^{n+\theta_{i+\frac{1}{2}}}) = 0. \quad (4.21)$$

Here  $\check{F} \check{U}_{i+\frac{1}{2}}^n$  accounts for the contribution of the nonlinear convective terms and the matrices  $\mathbf{R}^{\text{FV}}$  and  $\mathbf{L}^{\text{FV}}$  approximate the spatial derivative in  $x$ -direction. Note that in the equations above we use  $\Delta x$  and not  $\Delta x_s$ , which makes appear a factor of  $N_s = 2P + 1$  in the matrices  $\mathbf{R}^{\text{FV}}$  and  $\mathbf{L}^{\text{FV}}$ . For clarity, we give an example for the approximation of  $\Delta x \partial U / \partial x$  with  $P = 2$  and  $\theta_{i \pm \frac{1}{2}} = 1$ :

$$\begin{aligned} \mathbf{R}^{\text{FV}} \cdot \check{U}_{i+\frac{1}{2}}^{n+1} - \mathbf{L}^{\text{FV}} \cdot \check{U}_{i-\frac{1}{2}}^{n+1} = \\ 5(\check{U}_{i-\frac{1}{2},4}^{n+1} - \check{U}_{i-\frac{1}{2},3}^{n+1} \check{U}_{i-\frac{1}{2},5}^{n+1} - \check{U}_{i-\frac{1}{2},4}^{n+1} \check{U}_{i+\frac{1}{2},1}^{n+1} - \check{U}_{i-\frac{1}{2},5}^{n+1} \check{U}_{i+\frac{1}{2},2}^{n+1} - \check{U}_{i+\frac{1}{2},1}^{n+1} \check{U}_{i+\frac{1}{2},3}^{n+1} - \check{U}_{i+\frac{1}{2},2}^{n+1} \check{U}_{i+\frac{1}{2},4}^{n+1}). \end{aligned} \quad (4.22)$$

Similarly to the pure DG case, all explicit terms are collected in  $\tilde{\mathbf{G}}_{i+\frac{1}{2}}^{n+1}$  as

$$\tilde{\mathbf{G}}_{i+\frac{1}{2}}^{n+1} = \mathbf{F}\tilde{\mathbf{U}}_{i+\frac{1}{2}}^n - g \frac{\Delta t}{\Delta x} (1 - \theta_{i+\frac{1}{2}}) \tilde{\mathbf{H}}_{i+\frac{1}{2}}^n \cdot (\mathbf{R}^{\text{FV}} \cdot \tilde{\boldsymbol{\eta}}_{i+1}^n - \mathbf{L}^{\text{FV}} \cdot \tilde{\boldsymbol{\eta}}_i^n) \quad (4.23)$$

and finally we get

$$\tilde{\mathbf{U}}_{i+\frac{1}{2}}^{n+1} = \tilde{\mathbf{G}}_{i+\frac{1}{2}}^n - g \frac{\Delta t}{\Delta x} \theta_{i+\frac{1}{2}} \tilde{\mathbf{H}}_{i+\frac{1}{2}}^n \cdot (\mathbf{R}^{\text{FV}} \cdot \tilde{\boldsymbol{\eta}}_{i+1}^{n+1} - \mathbf{L}^{\text{FV}} \cdot \tilde{\boldsymbol{\eta}}_i^{n+1}). \quad (4.24)$$

From eq. (4.24) it is possible to observe that steady lake at rest solutions are exactly preserved also if the bathymetry is not constant; thus we can state that the sub-cell formulation of the semi-implicit finite volume method is well-balanced by construction.

Inserting eq. (4.24) into eq. (4.20) yields the following linear block three-diagonal system

$$\begin{aligned} \tilde{\boldsymbol{\eta}}_i^{n+1} + g \frac{\Delta t^2}{\Delta x^2} \left[ \theta_{i+\frac{1}{2}}^2 \mathbf{R}^{\text{FV}} \cdot \left( \tilde{\boldsymbol{\mathcal{H}}}_{i+\frac{1}{2}}^n \cdot \mathbf{R}^{\text{FV}} \cdot \tilde{\boldsymbol{\eta}}_{i+1}^{n+1} - \tilde{\boldsymbol{\mathcal{H}}}_{i+\frac{1}{2}}^n \cdot \mathbf{L}^{\text{FV}} \cdot \tilde{\boldsymbol{\eta}}_i^{n+1} \right) \right. \\ \left. - \theta_{i-\frac{1}{2}}^2 \mathbf{L}^{\text{FV}} \cdot \left( \tilde{\boldsymbol{\mathcal{H}}}_{i-\frac{1}{2}}^n \cdot \mathbf{R}^{\text{FV}} \cdot \tilde{\boldsymbol{\eta}}_i^{n+1} - \tilde{\mathbf{H}}_{i-\frac{1}{2}}^n \cdot \mathbf{L}^{\text{FV}} \cdot \tilde{\boldsymbol{\eta}}_{i-1}^{n+1} \right) \right] = \tilde{\mathbf{b}}_i^n, \end{aligned} \quad (4.25)$$

or more compactly

$$\mathcal{L}_i^{\text{FV}} \cdot \tilde{\boldsymbol{\eta}}_{i-1}^{n+1} + \mathcal{C}_i^{\text{FV}} \cdot \tilde{\boldsymbol{\eta}}_i^{n+1} + \mathcal{R}_i^{\text{FV}} \cdot \tilde{\boldsymbol{\eta}}_{i+1}^{n+1} = \tilde{\mathbf{b}}_i^n, \quad (4.26)$$

where  $\tilde{\boldsymbol{\mathcal{H}}}_{i\pm\frac{1}{2}}^n = \text{diag}(\tilde{\mathbf{H}}_{i\pm\frac{1}{2}}^n)$  and  $\tilde{\mathbf{b}}_i^n$  is the right hand side vector that contains the information at time  $t^n$

$$\begin{aligned} \tilde{\mathbf{b}}_i^n = \tilde{\boldsymbol{\eta}}_i^n - \frac{\Delta t}{\Delta x} (\theta_{i+\frac{1}{2}} \mathbf{R}^{\text{FV}} \cdot \tilde{\mathbf{G}}_{i+\frac{1}{2}}^n - \theta_{i-\frac{1}{2}} \mathbf{L}^{\text{FV}} \cdot \tilde{\mathbf{G}}_{i-\frac{1}{2}}^n) \\ - \frac{\Delta t}{\Delta x} ((1 - \theta_{i+\frac{1}{2}}) \mathbf{R}^{\text{FV}} \cdot \tilde{\mathbf{U}}_{i+\frac{1}{2}}^n - (1 - \theta_{i-\frac{1}{2}}) \mathbf{L}^{\text{FV}} \cdot \tilde{\mathbf{U}}_{i-\frac{1}{2}}^n). \end{aligned} \quad (4.27)$$

Here, it is easy to verify that the time integration of the momentum equation, done using the parameter  $\theta_{i+\frac{1}{2}}$ , allows to keep the symmetry of the linear system in eq. (4.25). System (4.26) can be solved either by the Thomas algorithm for block systems or by an iterative conjugate gradient method with a proper matrix-free formulation. Later on, the momentum equation (4.20) is applied in order to update the quantity  $\tilde{\mathbf{U}}_{i+\frac{1}{2}}^{n+1}$ . The water depth is computed by averaging the free surface from the main grid to the dual mesh as

$$\tilde{\mathbf{H}}_{i+\frac{1}{2}} = \tilde{\mathbf{h}}_{i+\frac{1}{2}} + (\mathbf{M}_{\mathbf{L}}^{\text{FV}} \cdot \tilde{\boldsymbol{\eta}}_i + \mathbf{M}_{\mathbf{R}}^{\text{FV}} \cdot \tilde{\boldsymbol{\eta}}_{i+1}). \quad (4.28)$$

Also in this case we compute the nonlinear convective terms on the main grid using the explicit finite volume discretization

$$\mathbf{F}\tilde{\mathbf{U}}_{i,s}^n = \tilde{\mathbf{U}}_{i,s}^n - \frac{\Delta t}{\Delta x} (f_{i,s+\frac{1}{2}}^n - f_{i,s-\frac{1}{2}}^n), \quad (4.29)$$

where the numerical flux  $f_{i,s+\frac{1}{2}}^n$  is again computed using a classical Rusanov method. Before this, we need to project  $u$  and  $U$  from the staggered grid to the main grid. For example  $\tilde{\mathbf{U}}_i$  is computed as

$$\tilde{\mathbf{U}}_i = (\mathbf{M}_{\mathbf{L}}^{\text{FV}} \cdot \tilde{\mathbf{U}}_{i-\frac{1}{2}} + \mathbf{M}_{\mathbf{R}}^{\text{FV}} \cdot \tilde{\mathbf{U}}_{i+\frac{1}{2}}). \quad (4.30)$$

Finally,  $\tilde{\mathbf{F}}\mathbf{U}_{i+\frac{1}{2}}$  is given by averaging back  $\tilde{\mathbf{F}}\mathbf{U}_i$  onto the staggered mesh using (4.28). The stability condition of the scheme is given by the maximum time step admissible for the explicit computation of the nonlinear convective terms, which in this case is  $\Delta t < \text{CFL} \frac{\Delta x_s}{2 \max(|u|)}$  with  $\text{CFL} < 1$ . Note that for  $\theta = 1$  the monotonicity in time of this finite volume method is ensured [75, 78].

#### 4.1.4 MOOD algorithm and detection criteria - 1D case

The limiting strategy adopted here is similar to the one proposed in [14, 59, 64] for the explicit case and is based on the MOOD paradigm introduced in [39, 47, 48, 107]. At each time step the numerical scheme is composed of two levels. In the first level, the unlimited DG scheme presented in section 4.1.2 is used to produce a so-called *candidate solution*  $\mathbf{Q}^{*,n+1} = (\eta^{*,n+1}, U^{*,n+1})$  at time  $t^{n+1}$ . Then, the troubled zones are identified using physical and numerical admissibility criteria. In troubled cells, the more robust staggered semi-implicit finite volume subcell scheme introduced in the previous section 4.1.3 is used, while those cells which are not troubled are run with the unlimited DG method. Therefore, in the second level of the algorithm, which starts again from a valid solution at time  $t^n$ , a *mixed* scheme is employed, with the more robust subcell finite volume method in troubled cells and the unlimited DG method in the other ones. For this purpose, the linear system for the free surface has to be assembled and solved again. In the case where no troubled cells have been detected after the first level of the algorithm, the scheme directly proceeds with the next time step without activating the finite volume subcell limiter and without having to come back to time  $t^n$  in the second level.

##### 4.1.4.1 Data representation, projection and reconstruction

If  $\hat{q}_i^n$  denotes the set of degrees of freedom used for the piecewise polynomial data representation of a generic quantity  $q$  of the numerical solution given by the DG scheme in cell  $T_i$  at time  $t^n$ , then the set of piecewise constant cell averages in the subcells  $T_{i,s}$  is denoted by  $\tilde{q}_i^n$ . Supposing that the coefficients  $\hat{q}_{i,l}^n$  are known, then  $\tilde{q}_{i,s}^n$  is computed using the following  $L_2$  projection

$$\tilde{q}_{i,s}^n = \frac{1}{|T_{i,s}|} \int_{T_{i,s}} \phi_l(x) dx \hat{q}_{i,l}^n, \quad \forall T_{i,s} \in T_i, \quad (4.31)$$

which defines the projection operator  $\mathcal{P}$  so that  $\tilde{q}_i^n = \mathcal{P} \cdot \hat{q}_i^n$ . On the contrary, we can gather back the piecewise constant sub-cell averages  $\tilde{q}_{i,s}^n$  into the coefficients  $\hat{q}_{i,l}^n$  of a high order DG polynomial by solving again (4.31), but which now leads to an *over-determined* problem, since that there are  $2P+1$  equations for  $N+1$  unknowns. In order to solve the problem, we use a *constrained least-squares* approach (see [57, 97]), where the linear constraint is integral conservation on the big cell  $T_i$  (see [64]), i.e.

$$\sum_j |T_{i,s}| \tilde{q}_{i,s}^n = \int_{T_i} \phi_l(x) dx \hat{q}_{i,l}^n. \quad (4.32)$$

This last procedure can also be expressed in a matrix-vector notation as  $\tilde{q}_i^n = \mathcal{W} \cdot \hat{q}_i^n$ , with  $\mathcal{W} \cdot \mathcal{P} = \mathcal{I}$ , where  $\mathcal{I}$  denotes the identity matrix. For a schematic representation of the projection and reconstruction process, see Figure 4.3. Moreover, as we will see later, in some cases we also

need to reconstruct a polynomial using only the  $P + 1$  piecewise constant sub-cell data at the left part,  $T_i^L = \bigcup_{s=1}^{P+1} T_{i,s}$  (see 4.4), or the right,  $T_i^R = \bigcup_{s=P+1}^{2P+1} T_{i,s}$  (see 4.5). So we introduce two additional operators  $\mathcal{W}_\mathcal{L}$  and  $\mathcal{W}_\mathcal{R}$  that are computed using (4.31) only on the left and right half of cell  $T_i$ , but without any least squares procedure, since now the number of equations is equal to the number of unknowns.

#### 4.1.4.2 Detection criteria

Here we describe the physical and numerical admissibility criteria that are applied to all quantities on the *main grid*. For those quantities which are located on the staggered dual grid, we use the algorithm given in eq. (4.15) to project them also onto the main grid. Once the candidate solution given by the unlimited staggered DG scheme has been obtained, we first require that the numerical solution must satisfy some *physical admissibility criteria* (PAD). For the shallow water equations we require that the water depth  $H$  is non-negative, hence checking if

$$\hat{H}_{i,l}^{*,n+1} \geq 0. \quad (4.33)$$

Then, according to [14, 59, 64], for the *numerical admissibility criteria* (NAD) we use a relaxed version of the discrete maximum principle (DMP) for a generic quantity  $q$  as

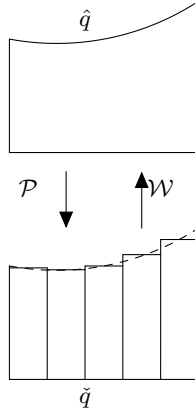
$$\min_{\forall T_j \in \mathcal{V}_i} (\hat{q}_{j,l}^n) - \delta \leq \hat{q}_{i,l}^{*,n+1} \leq \max_{\forall T_j \in \mathcal{V}_i} (\hat{q}_{j,l}^n) + \delta, \quad (4.34)$$

where  $\mathcal{V}_i = \{T_{i-1}, T_i, T_{i+1}\}$  is the set of neighbour cells of  $T_i$ , which will become the set of Voronoi neighbours in the two-dimensional case. In our implementation we check the DMP on the discrete free surface elevation and the discharge. Moreover, numerical experiments have shown that it may be useful to project the numerical quantities from the main grid to the dual mesh and then project them back onto the main grid. In this way additional non-physical oscillations are created near discontinuities, which allows an easier detection of troubled cells. In particular, this method also allows a robust detection of troubled cells in the discrete representation of the initial condition. In equation (4.34),  $\delta$  is a tolerance that allows some small overshoots and thus slightly *relaxes* the discrete maximum principle. It is computed as follows:

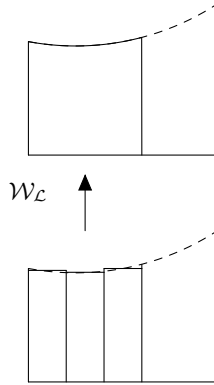
$$\delta = \max[\delta_0, \epsilon (\max_{\forall T_j \in \mathcal{V}_i} (\hat{q}_{j,l}^n) - \min_{\forall T_j \in \mathcal{V}_i} (\hat{q}_{j,l}^n))], \quad (4.35)$$

where  $\delta_0$  is a parameter taken in the interval  $[10^{-4}, 10^{-3}]$  and typically  $\epsilon = 5 \cdot 10^{-4}$ . If a cell does not pass one of these *a posteriori* detection criteria applied to the candidate solution at time  $t^{n+1}$  it is flagged as troubled, which means that the solution will be recomputed in this cell using the finite volume scheme presented in Section 4.1.3. A cell  $T_i$  is flagged as troubled by setting an indicator parameter  $\beta_i = 1$ , while untroubled cells are flagged with  $\beta_i = 0$ .

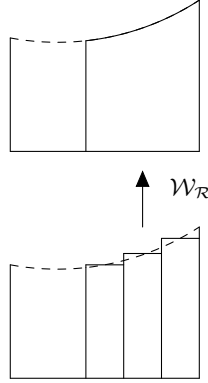
In order to extend the troubled cell indicator properly to the dual mesh, if the cell  $T_i$  is troubled then we also flag both dual control volumes  $T_{i-\frac{1}{2}}$  and  $T_{i+\frac{1}{2}}$  as troubled, consequently setting  $\beta_{i-\frac{1}{2}} = 1$  and  $\beta_{i+\frac{1}{2}} = 1$ , see Figure 4.6 for a schematic representation of the data in this case. Moreover, in the first MOOD iteration the parameter  $\theta = \theta_{DG}$  is chosen as a constant value  $\theta_{DG}$  in all the cells of the dual grid and  $\theta_{DG}$  is chosen very close to but larger than  $1/2$  in order to reduce the numerical viscosity as much as possible. In the second MOOD step, the value of the parameter  $\theta$  is changed to  $\theta = 1$  in the troubled cells, where the subcell finite volume method is used. Note that, if these zones are properly detected, the monotonicity of the global method is assured.



**Figure 4.3:** Projection and reconstruction operators - Operators  $\mathcal{P}$  and  $\mathcal{W}$



**Figure 4.4:** Projection and reconstruction operators - Operator  $\mathcal{W}_{\mathcal{L}}$



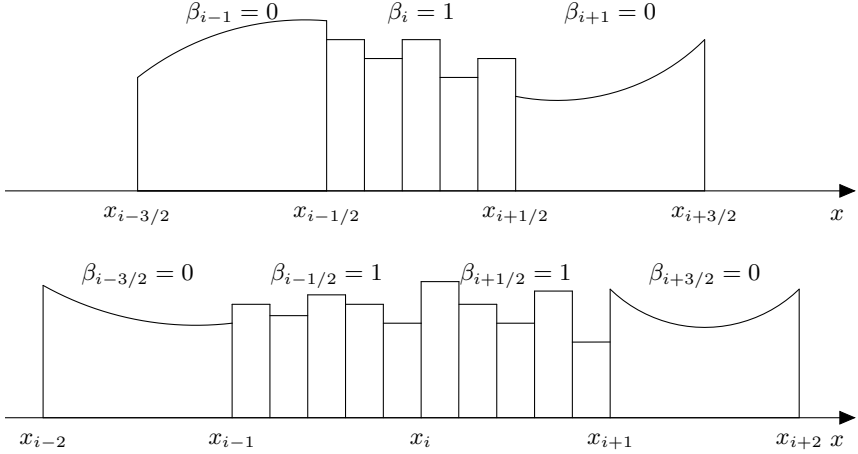
**Figure 4.5:** Projection and reconstruction operators - Operator  $\mathcal{W}_{\mathcal{R}}$

#### 4.1.5 Sub-cell limiting of the semi-implicit DG scheme for the 1D shallow water equations

We start from the discretization of the continuity equation and for the general algorithm including the limiter we have to consider four cases. The first one is a non-troubled cell on the main grid,  $\beta_i = 0$ , and the corresponding cells on the dual grid are not limited, i.e.  $\beta_{i-\frac{1}{2}} = \beta_{i+\frac{1}{2}} = 0$ . This case corresponds to the pure DG scheme exposed in subsection 4.1.2. On the other hand, there is the opposite situation when all control volumes  $T_i, T_{i-\frac{1}{2}}, T_{i+\frac{1}{2}}$  are occupied by troubled cells. Consequently, the discrete continuity equation is the one presented in eq. (4.20). Finally, if  $T_i$  is not troubled then still one of the neighbour elements on the dual mesh can be troubled and the other not. Consequently, we distinguish the two possible cases for  $\beta_i = 0$  when  $\beta_{i-\frac{1}{2}} = 0$  and  $\beta_{i+\frac{1}{2}} = 1$  (see Fig. 4.7) and the other one when  $\beta_{i-\frac{1}{2}} = 1$  and  $\beta_{i+\frac{1}{2}} = 0$  (see Fig. 4.8). These two situations are governed respectively by eq. (4.36) and by eq. (4.37) below:

$$\mathbf{M} \cdot \left( \frac{\mathbf{n}_i^{n+1} - \hat{\mathbf{n}}_i^n}{\Delta t} \right) + \frac{\mathbf{R}_J^{\text{DG}} \cdot \tilde{\mathbf{U}}_{i+\frac{1}{2}}^{n+\theta} - \mathbf{L}_J^{\text{Lim}} \cdot \tilde{\mathbf{U}}_{i-\frac{1}{2}}^{n-\theta}}{\Delta x} = 0 \quad \text{with} \quad \mathbf{L}_J^{\text{Lim}} = \mathbf{L}_J^{\text{DG}} \cdot \mathcal{W}_{\mathcal{R}} \quad (4.36)$$



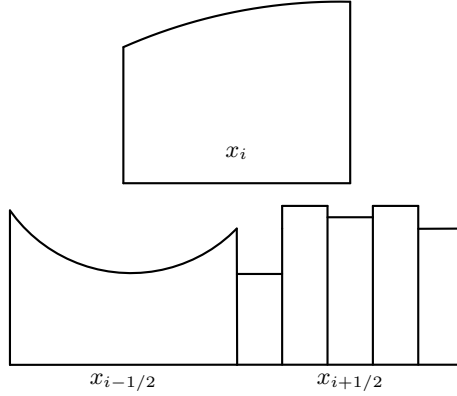


**Figure 4.6:** Staggered grids for the semi-implicit staggered DG scheme with sub-cell limiter active in cell  $T_i$ . Main grid used for the free surface (top) and staggered dual mesh for the velocity (bottom). If a cell  $T_i$  on the main grid is flagged as troubled, then also the two overlapping staggered velocity control volumes  $T_{i \pm \frac{1}{2}}$  are flagged as troubled. The data representation in troubled cells is changed from high order polynomials to piecewise constant subcell averages.

$$\mathbf{M} \cdot \left( \frac{\hat{\eta}_i^{n+1} - \hat{\eta}_i^n}{\Delta t} \right) + \frac{\mathbf{R}_J^{\text{Lim}} \cdot \tilde{\mathbf{U}}_{i+\frac{1}{2}}^{n+\theta} - \mathbf{L}_J^{\text{DG}} \cdot \hat{\mathbf{U}}_{i-\frac{1}{2}}^{n+\theta}}{\Delta x} = 0 \quad \text{with} \quad \mathbf{R}_J^{\text{Lim}} = \mathbf{R}_J^{\text{DG}} \cdot \mathcal{W}_{\mathcal{L}}. \quad (4.37)$$

Note that the structure of these two equations is very similar to the one of the discrete continuity equation in the pure DG case, eq. (4.11). In particular, looking at Fig. 4.7, the operator  $\mathcal{W}_{\mathcal{L}}$  reconstructs the polynomial data using the first  $P+1$  sub-cell finite volume data of the cell  $T_{i+\frac{1}{2}}$ . A similar consideration is valid for Fig. 4.8, where  $\mathcal{W}_{\mathcal{R}}$  builds the polynomial information based on the last  $P+1$  sub-cell finite volume data of the cell  $T_{i-\frac{1}{2}}$ .

Also for the general case of the momentum equation there are four possible combinations. The



**Figure 4.7:** Control volumes for the continuity equation for the staggered semi-implicit DG scheme in the limited case using  $P = 2$ , i.e.  $N_s = 2P + 1 = 5$ . DG method used in  $T_i$  and  $T_{i-\frac{1}{2}}$ ; finite volumes in  $T_{i+\frac{1}{2}}$

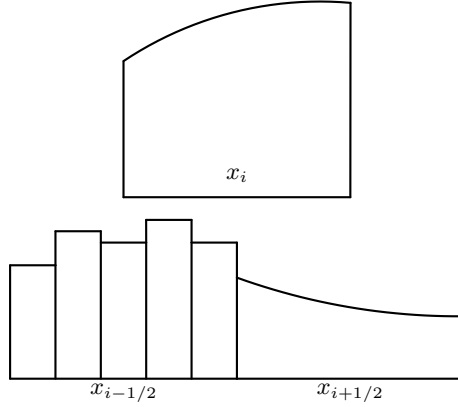
first one is the pure DG case where the discrete momentum equation reduces to eq. (4.11). On the contrary, when  $\beta_i = \beta_{i+1/2} = \beta_{i+1} = 1$  the sub-cell finite volume scheme is locally applied in all cells, hence the degrees of freedom of  $\tilde{U}_{i+\frac{1}{2}}$  are given by eq. (4.24). The interesting cases are when  $T_{i+\frac{1}{2}}$  is limited and only one of the control volumes  $T_i$  and  $T_{i+1}$  is a troubled cell. Hence, when  $\beta_i = 0$  and  $\beta_{i+1} = 1$  (see Fig. 4.9) the numerical expression of the discrete momentum equation reads

$$\tilde{U}_{i+\frac{1}{2}}^{n+1} = \mathbf{F}\tilde{U}_{i+\frac{1}{2}}^n - g \frac{\Delta t}{\Delta x} \tilde{H}_{i+\frac{1}{2}}^n \cdot (\mathbf{R}_{\mathbf{u}}^{\text{Lim}} \cdot \hat{\boldsymbol{\eta}}_{i+1}^{n+\theta_{i+\frac{1}{2}}} - \mathbf{L}^{\text{FV}} \cdot \tilde{\boldsymbol{\eta}}_i^{n+\theta_{i+\frac{1}{2}}}) \quad \text{with} \quad \mathbf{R}_{\mathbf{u}}^{\text{Lim}} = \mathbf{R}^{\text{FV}} \cdot \mathcal{P}, \quad (4.38)$$

and when  $\beta_i = 1$  and  $\beta_{i+1} = 0$  (see Fig. 4.10) we get

$$\tilde{U}_{i+\frac{1}{2}}^{n+1} = \mathbf{F}\tilde{U}_{i+\frac{1}{2}}^n - g \frac{\Delta t}{\Delta x} \tilde{H}_{i+\frac{1}{2}}^n \cdot (\mathbf{R}^{\text{FV}} \cdot \tilde{\boldsymbol{\eta}}_{i+1}^{n+\theta_{i+\frac{1}{2}}} - \mathbf{L}_{\mathbf{u}}^{\text{Lim}} \cdot \hat{\boldsymbol{\eta}}_i^{n+\theta_{i+\frac{1}{2}}}) \quad \text{with} \quad \mathbf{L}_{\mathbf{u}}^{\text{Lim}} = \mathbf{L}^{\text{FV}} \cdot \mathcal{P}. \quad (4.39)$$

Note that, in eq. (4.38), the product  $\mathbf{R}_{\mathbf{u}}^{\text{Lim}} \cdot \hat{\boldsymbol{\eta}}_{i+1}^{n+\theta_{i+\frac{1}{2}}} = \mathbf{R}^{\text{FV}} \cdot \tilde{\boldsymbol{\eta}}_{i+1}^{n+\theta_{i+\frac{1}{2}}}$  represents the multiplication of the matrix  $\mathbf{R}^{\text{FV}}$  with the vector of the finite volume degrees of freedom  $\tilde{\boldsymbol{\eta}}_{i+1}^{n+\theta_{i+\frac{1}{2}}} =$



**Figure 4.8:** Control volumes for the continuity equation for the staggered semi-implicit DG scheme in the limited case using  $P = 2$ , i.e.  $N_s = 2P + 1 = 5$ . DG method used in  $T_i$  and  $T_{i+\frac{1}{2}}$ ; finite volumes in  $T_{i-\frac{1}{2}}$ .

$\mathcal{P}\hat{\eta}_{i+1}^{n+\theta_{i+\frac{1}{2}}}$  obtained from projection of the DG degrees of freedom onto the set of piecewise constant subcell averages. The same consideration is valid also for the product  $\mathbf{L}_{\mathbf{u}}^{\text{Lim}} \cdot \hat{\eta}_i^{n+\theta_{i+\frac{1}{2}}} = \mathbf{L}^{\text{FV}} \cdot \tilde{\eta}_i^{n+\theta_{i+\frac{1}{2}}}$  in eq. (4.39). Hence, equations (4.38) and (4.39) can be seen as special cases of eq. (4.24) and thus the C-property is valid also in the limited case.

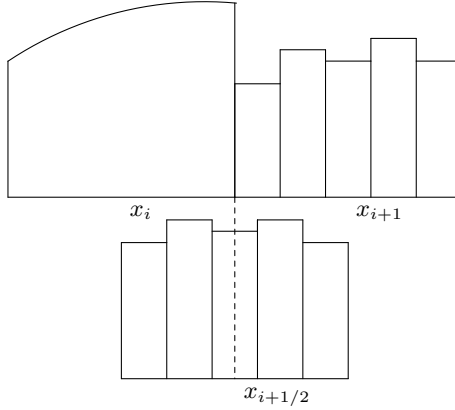
Again, combining the continuity equation with the momentum equation yields the following block three-diagonal system, where the generic degrees of freedom of the free surface are now denoted by the symbol  $\tilde{\eta}_i^{n+1}$ , representing either finite volume data or DG data, depending on the limiter flag  $\beta_i$ ,

$$\mathcal{L}_i^{\text{Lim}} \cdot \tilde{\eta}_{i-1}^{n+1} + \mathcal{C}_i^{\text{Lim}} \cdot \tilde{\eta}_i^{n+1} + \mathcal{R}_i^{\text{Lim}} \cdot \tilde{\eta}_{i+1}^{n+1} = \tilde{\mathbf{b}}_i^n. \quad (4.40)$$

The solution can be efficiently obtained by the Thomas algorithm for block three-diagonal systems. For the sake of clarity, we split the block  $\mathcal{C}_i^{\text{Lim}}$  into two contributions:

$$\mathcal{C}_i^{\text{Lim}} = \mathcal{C}_i^{0,\text{Lim}} + \mathcal{C}_i^{x,\text{Lim}}, \quad (4.41)$$

where  $\mathcal{C}_i^{0,\text{Lim}}$  is the mass matrix  $\mathbf{M}$  if cell  $i$  is not limited or the identity matrix  $\mathbf{I}$  if it is limited, i.e.

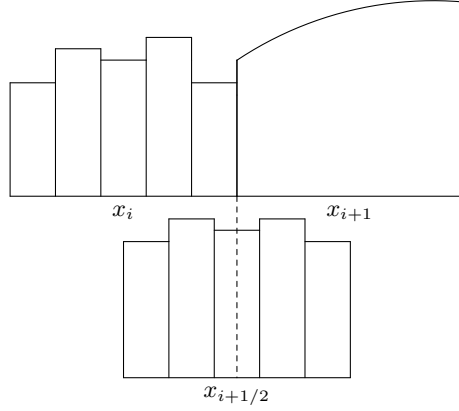


**Figure 4.9:** Control volumes for the momentum equation for the staggered semi-implicit DG scheme in the limited case using  $P = 2$  ( $N_s = 5$ ). DG in  $T_i$ ; FV in  $T_{i+1}$  and in  $T_{i+\frac{1}{2}}$

$$\mathbf{C}_i^{0,\text{Lim}} = \begin{cases} \mathbf{M} & \text{if } \beta_i = 0, \\ \mathbf{I} & \text{if } \beta_i = 1. \end{cases} \quad (4.42)$$

Then, depending on the distribution of the limiter flag  $\beta$ , the blocks  $\mathcal{L}_i^{\text{Lim}}$ ,  $\mathcal{C}_i^{x,\text{Lim}}$  and  $\mathcal{R}_i^{\text{Lim}}$  are computed as follows

$$\mathcal{L}_i^{\text{Lim}} = -g \frac{\Delta t^2}{\Delta x^2} \theta^2_{i-\frac{1}{2}} \begin{cases} \mathbf{L}_j^{\text{DG}} \cdot \mathbf{M}^{-1} \cdot \mathbf{L}_u^{\text{DG}} \cdot \hat{\mathbf{H}}_{i-\frac{1}{2}}^n & \text{if } \beta_{i-1} = \beta_i = 0, \\ \mathbf{L}^{\text{FV}} \cdot \hat{\mathcal{H}}_{i-\frac{1}{2}}^n \cdot \mathbf{L}^{\text{FV}} & \text{if } \beta_{i-1} = \beta_i = 1, \\ \mathbf{L}^{\text{FV}} \cdot \hat{\mathcal{H}}_{i-\frac{1}{2}}^n \cdot \mathbf{L}_u^{\text{Lim}} & \text{if } \beta_{i-1} = 1, \beta_i = 0, \\ \mathbf{L}_j^{\text{Lim}} \cdot \hat{\mathcal{H}}_{i-\frac{1}{2}}^n \cdot \mathbf{L}^{\text{FV}} & \text{if } \beta_{i-1} = 0, \beta_i = 1, \end{cases} \quad (4.43)$$



**Figure 4.10:** Control volumes for the momentum equation for the staggered semi-implicit DG scheme in the limited case using  $P = 2$  ( $N_s = 5$ ). DG in  $T_{i+1}$ ; FV in  $T_i$  and in  $T_{i+\frac{1}{2}}$

$$C_i^{x,\text{Lim}} = g \frac{\Delta t^2}{\Delta x^2} \begin{cases} \theta_{i-\frac{1}{2}}^2 \mathbf{L}_J^{\text{DG}} \cdot \mathbf{M}^{-1} \cdot \mathbf{R}_u^{\text{DG}} \cdot \hat{\mathbf{H}}_{i-\frac{1}{2}}^n + \theta_{i+\frac{1}{2}}^2 \mathbf{R}_J^{\text{DG}} \cdot \mathbf{M}^{-1} \cdot \mathbf{L}_u^{\text{DG}} \cdot \hat{\mathbf{H}}_{i+\frac{1}{2}}^n \\ \text{if } \beta_i = \beta_{i-\frac{1}{2}} = \beta_{i+\frac{1}{2}} = 0, \\ \theta_{i-\frac{1}{2}}^2 \mathbf{L}_J^{\text{FV}} \cdot \tilde{\mathcal{H}}_{i-\frac{1}{2}}^n \cdot \frac{1}{2} \cdot \mathbf{R}^{\text{FV}} + \theta_{i+\frac{1}{2}}^2 \mathbf{R}_J^{\text{FV}} \cdot \tilde{\mathcal{H}}_{i+\frac{1}{2}}^n \cdot \frac{1}{2} \cdot \mathbf{L}^{\text{FV}} \\ \text{if } \beta_i = \beta_{i-\frac{1}{2}} = \beta_{i+\frac{1}{2}} = 1, \\ \theta_{i-\frac{1}{2}}^2 \mathbf{L}_J^{\text{DG}} \cdot \mathbf{M}^{-1} \cdot \mathbf{R}_u^{\text{DG}} \cdot \hat{\mathbf{H}}_{i-\frac{1}{2}}^n + \theta_{i+\frac{1}{2}}^2 \mathbf{R}_J^{\text{Lim}} \cdot \tilde{\mathcal{H}}_{i+\frac{1}{2}}^n \cdot \mathbf{L}_u^{\text{Lim}} \\ \text{if } \beta_i = \beta_{i-\frac{1}{2}} = 0, \beta_{i+\frac{1}{2}} = 1, \\ \theta_{i-\frac{1}{2}}^2 \mathbf{L}_J^{\text{Lim}} \cdot \tilde{\mathcal{H}}_{i-\frac{1}{2}}^n \cdot \mathbf{R}_u^{\text{Lim}} + \theta_{i+\frac{1}{2}}^2 \mathbf{R}_J^{\text{DG}} \cdot \mathbf{M}^{-1} \cdot \mathbf{L}_u^{\text{DG}} \cdot \hat{\mathbf{H}}_{i+\frac{1}{2}}^n \\ \text{if } \beta_i = \beta_{i+\frac{1}{2}} = 0, \beta_{i-\frac{1}{2}} = 1, \end{cases} \quad (4.44)$$

$$\mathcal{R}_i^{\text{Lim}} = -g \frac{\Delta t^2}{\Delta x^2} \theta_{i+\frac{1}{2}}^2 \begin{cases} \mathbf{R}_J^{\text{DG}} \cdot \mathbf{M}^{-1} \cdot \mathbf{R}_u^{\text{DG}} \cdot \hat{\mathbf{H}}_{i+\frac{1}{2}}^n & \text{if } \beta_{i+1} = \beta_i = 0, \\ \mathbf{R}^{\text{FV}} \cdot \tilde{\mathcal{H}}_{i+\frac{1}{2}}^n \cdot \mathbf{R}^{\text{FV}} & \text{if } \beta_{i+1} = \beta_i = 1, \\ \mathbf{R}^{\text{FV}} \cdot \tilde{\mathcal{H}}_{i+\frac{1}{2}}^n \cdot \mathbf{R}_u^{\text{Lim}} & \text{if } \beta_{i+1} = 1, \beta_i = 0, \\ \mathbf{R}_J^{\text{Lim}} \cdot \tilde{\mathcal{H}}_{i+\frac{1}{2}}^n \cdot \mathbf{R}^{\text{FV}} & \text{if } \beta_{i+1} = 0, \beta_i = 1. \end{cases} \quad (4.45)$$

Note that the first cases correspond to the pure DG method, the second cases to the sub-cell finite volume method and the other two cases are those for the limited situations. Also for the term on the right hand side it is convenient to write the two parts as follows

$$\tilde{\mathbf{b}}_i^n = \tilde{\mathbf{b}}_i^{0,n} + \tilde{\mathbf{b}}_i^{x,n} \quad (4.46)$$

with

$$\tilde{\mathbf{b}}_i^{0,n} = \begin{cases} \mathbf{M} \cdot \hat{\boldsymbol{\eta}}_i^n & \text{if } \beta_i = 0, \\ \tilde{\boldsymbol{\eta}}_i^n & \text{if } \beta_i = 1, \end{cases} \quad (4.47)$$

and with

$$\tilde{\mathbf{b}}_i^{x,n} = -\frac{\Delta t}{\Delta x} \begin{cases} (\theta_{i+\frac{1}{2}} \mathbf{R}_J^{\text{DG}} \cdot \hat{\mathbf{G}}_{i+\frac{1}{2}}^n - \theta_{i-\frac{1}{2}} \mathbf{L}_J^{\text{DG}} \cdot \hat{\mathbf{G}}_{i-\frac{1}{2}}^n) \\ + ((1 - \theta_{i+\frac{1}{2}}) \mathbf{R}_J^{\text{DG}} \cdot \tilde{\mathbf{U}}_{i+\frac{1}{2}}^n - (1 - \theta_{i-\frac{1}{2}}) \mathbf{L}_J^{\text{DG}} \cdot \tilde{\mathbf{U}}_{i-\frac{1}{2}}^n) \\ \text{if } \beta_i = \beta_{i-\frac{1}{2}} = \beta_{i+\frac{1}{2}} = 0, \\ (\theta_{i+\frac{1}{2}} \mathbf{R}^{\text{FV}} \cdot \hat{\mathbf{G}}_{i+\frac{1}{2}}^n - \theta_{i-\frac{1}{2}} \mathbf{L}^{\text{FV}} \cdot \check{\mathbf{G}}_{i-\frac{1}{2}}^n) \\ + ((1 - \theta_{i+\frac{1}{2}}) \mathbf{R}^{\text{FV}} \cdot \check{\mathbf{U}}_{i+\frac{1}{2}}^n - (1 - \theta_{i-\frac{1}{2}}) \mathbf{L}^{\text{FV}} \cdot \check{\mathbf{U}}_{i-\frac{1}{2}}^n) \\ \text{if } \beta_i = \beta_{i-\frac{1}{2}} = \beta_{i+\frac{1}{2}} = 1, \\ (\theta_{i+\frac{1}{2}} \mathbf{R}_J^{\text{Lim}} \cdot \hat{\mathbf{G}}_{i+\frac{1}{2}}^n - \theta_{i-\frac{1}{2}} \mathbf{L}_J^{\text{DG}} \cdot \hat{\mathbf{G}}_{i-\frac{1}{2}}^n) \\ + ((1 - \theta_{i+\frac{1}{2}}) \mathbf{R}_J^{\text{Lim}} \cdot \tilde{\mathbf{U}}_{i+\frac{1}{2}}^n - (1 - \theta_{i-\frac{1}{2}}) \mathbf{L}_J^{\text{DG}} \cdot \tilde{\mathbf{U}}_{i-\frac{1}{2}}^n) \\ \text{if } \beta_i = \beta_{i-\frac{1}{2}} = 0, \beta_{i+\frac{1}{2}} = 1, \\ (\theta_{i+\frac{1}{2}} \mathbf{R}_J^{\text{DG}} \cdot \hat{\mathbf{G}}_{i+\frac{1}{2}}^n - \theta_{i-\frac{1}{2}} \mathbf{L}_J^{\text{Lim}} \cdot \check{\mathbf{G}}_{i-\frac{1}{2}}^n) \\ + ((1 - \theta_{i+\frac{1}{2}}) \mathbf{R}_J^{\text{DG}} \cdot \tilde{\mathbf{U}}_{i+\frac{1}{2}}^n - (1 - \theta_{i-\frac{1}{2}}) \mathbf{L}_J^{\text{Lim}} \cdot \check{\mathbf{U}}_{i-\frac{1}{2}}^n) \\ \text{if } \beta_i = \beta_{i+\frac{1}{2}} = 0, \beta_{i-\frac{1}{2}} = 1. \end{cases} \quad (4.48)$$

When the degrees of freedom of the free surface are known, the variable  $\tilde{\mathbf{U}}_{i+\frac{1}{2}}^{n+1}$  is easily computed using the discrete momentum equations (4.11), (4.24), (4.38) or (4.39), depending on  $\beta$ . Then, the algorithm for the extrapolation of the water depth  $\tilde{\mathbf{H}}_{i+\frac{1}{2}}^{n+1}$  on the dual grid reads as follows

$$\tilde{\mathbf{H}}_{i+\frac{1}{2}} = \begin{cases} \hat{h}_{i+\frac{1}{2}} + \mathbf{M}^{-1} \cdot (\mathbf{M}_L^{\text{DG}} \cdot \hat{\boldsymbol{\eta}}_i + \mathbf{M}_R^{\text{DG}} \cdot \hat{\boldsymbol{\eta}}_{i+1}) & \text{if } \beta_{i+\frac{1}{2}} = \beta_i = \beta_{i+1} = 0, \\ \hat{h}_{i+\frac{1}{2}} + (\mathbf{M}_L^{\text{FV}} \cdot \tilde{\boldsymbol{\eta}}_i + \mathbf{M}_R^{\text{FV}} \cdot \tilde{\boldsymbol{\eta}}_{i+1}) & \text{if } \beta_{i+\frac{1}{2}} = \beta_i = \beta_{i+1} = 1, \\ \check{h}_{i+\frac{1}{2}} + (\mathbf{M}_L^{\text{FV}} \cdot \mathcal{P} \cdot \hat{\boldsymbol{\eta}}_i + \mathbf{M}_R^{\text{FV}} \cdot \tilde{\boldsymbol{\eta}}_{i+1}) & \text{if } \beta_{i+\frac{1}{2}} = \beta_{i+1} = 1, \beta_i = 0, \\ \check{h}_{i+\frac{1}{2}} + (\mathbf{M}_L^{\text{FV}} \cdot \tilde{\boldsymbol{\eta}}_i + \mathbf{M}_R^{\text{FV}} \cdot \mathcal{P} \cdot \hat{\boldsymbol{\eta}}_{i+1}) & \text{if } \beta_{i+\frac{1}{2}} = \beta_i = 1, \beta_{i+1} = 0. \end{cases} \quad (4.49)$$

Finally, we compute the nonlinear contribution of the convective terms  $\tilde{\mathbf{F}}\tilde{\mathbf{U}}_i$  and also in the limited case it is evaluated on the main grid. First, in order to project the discharge  $\tilde{\mathbf{U}}_{i+\frac{1}{2}}$  on the primal grid we use the following procedure

$$\tilde{\mathbf{U}}_i = \begin{cases} \mathbf{M}^{-1} \cdot (\mathbf{M}_L^{\text{DG}} \cdot \tilde{\mathbf{U}}_{i-\frac{1}{2}} + \mathbf{M}_R^{\text{DG}} \cdot \tilde{\mathbf{U}}_{i+\frac{1}{2}}) & \text{if } \beta_i = \beta_{i-\frac{1}{2}} = \beta_{i+\frac{1}{2}} = 0, \\ \mathbf{M}_L^{\text{FV}} \cdot \tilde{\mathbf{U}}_{i-\frac{1}{2}} + \mathbf{M}_R^{\text{FV}} \cdot \tilde{\mathbf{U}}_{i+\frac{1}{2}} & \text{if } \beta_i = \beta_{i-\frac{1}{2}} = \beta_{i+\frac{1}{2}} = 1, \\ \mathbf{M}^{-1} \cdot (\mathbf{M}_L^{\text{DG}} \cdot \tilde{\mathbf{U}}_{i-\frac{1}{2}} + \mathbf{M}_R^{\text{DG}} \cdot \mathcal{W}_L \cdot \tilde{\mathbf{U}}_{i+\frac{1}{2}}) & \text{if } \beta_i = \beta_{i-\frac{1}{2}} = 0, \beta_{i+\frac{1}{2}} = 1, \\ \mathbf{M}^{-1} \cdot (\mathbf{M}_L^{\text{DG}} \cdot \mathcal{W}_R \cdot \tilde{\mathbf{U}}_{i-\frac{1}{2}} + \mathbf{M}_R^{\text{DG}} \cdot \tilde{\mathbf{U}}_{i+\frac{1}{2}}) & \text{if } \beta_i = \beta_{i+\frac{1}{2}} = 0, \beta_{i-\frac{1}{2}} = 1, \end{cases} \quad (4.50)$$

which is applied also to the velocity  $\tilde{\mathbf{u}}_{i+\frac{1}{2}}$  in order to get  $\tilde{\mathbf{u}}_i$ . Then, if a cell is not troubled,  $\beta_i = 0$ , we use the explicit upwind algorithm used for the pure DG scheme explained at the end of section 4.1.2. Note that this explicit DG method is applied only in the first stage of the MOOD algorithm, while in the second part the nonlinear convective terms are recomputed only for the troubled cells using the classical explicit upwind finite volume method on the sub-grid shown in section 4.1.3. Note that this approach is essentially the one-dimensional case of the sub-cell limiting of the DG scheme proposed for the first time in [64]. Moreover, it is easy to check that due to the choice  $N_s = 2P + 1$  the CFL stability condition on the maximum admissible time step size  $\Delta t$  for the explicit DG scheme and for the explicit subgrid finite volume scheme are the same.

#### 4.1.6 A high-resolution wetting and drying algorithm for high order staggered semi-implicit DG in 1D

In this subsection, we use the numerical method introduced before in order to solve problems that include wet-dry fronts. The key idea is to activate the sub-cell limiter in the partially wet regions where the pure DG method would generated negative water depths. The wetting and drying approach that we follow is the strategy introduced by Casulli in [28]. At first, in the continuity equation of system (4.1), the time derivative of the free surface elevation is replaced by the time derivative of the water depth

$$\frac{\partial H}{\partial t} + \frac{\partial U}{\partial x} = 0, \quad (4.51)$$

since the bottom topography does not change in time. The new continuity equation is numerically approximated as follows

$$\frac{\tilde{\mathcal{H}}(\tilde{\eta}_i^{n+1}) - \tilde{\mathcal{H}}(\tilde{\eta}_i^n)}{\Delta t} + \frac{\Delta \tilde{U}_i^{n+\theta_i}}{\Delta x} = 0, \quad (4.52)$$

where the spatial derivative is discretized in the same way shown in section 4.1.5, while in the discrete time derivative now we have the nonlinear term

$$\tilde{\mathcal{H}}(\tilde{\eta}_i) = \begin{cases} \mathbf{M} \cdot \hat{\mathcal{H}}(\hat{\eta}_i) & \text{if } \beta_i = 0, \\ \tilde{\mathcal{H}}(\tilde{\eta}_i) & \text{if } \beta_i = 1. \end{cases} \quad (4.53)$$

In particular, for the degrees of freedom in the cells where the limiter is not activated we use

$$\tilde{\mathcal{H}}(\tilde{\eta}_{i,ii}) = \max[0, \tilde{h}_{i,ii} + \tilde{\eta}_{i,ii}] \quad (4.54)$$

while in the sub-cells of the limited control volumes we use the subgrid approach of Casulli and Stelling [35]. Hence, for every sub control volume  $\tilde{\mathcal{H}}(\tilde{\eta}_{i,s})$  is given by

$$\tilde{\mathcal{H}}(\tilde{\eta}_{i,s}) = \frac{1}{\Delta x_s} \int_{T_{i,s}} H(x, \eta) dx, \quad \text{with } H(x, \eta) = \max[0, h(x) + \eta]. \quad (4.55)$$

Finally, we get a mildly non-linear system that reads

$$\tilde{\mathcal{H}}(\tilde{\eta}_i^{n+1}) + \mathcal{L}_i^{\text{Lim}} \cdot \tilde{\eta}_{i-1}^{n+1} + \mathcal{C}_i^{x, \text{Lim}} \cdot \tilde{\eta}_i^{n+1} + \mathcal{R}_i^{\text{Lim}} \cdot \tilde{\eta}_{i+1}^{n+1} = \tilde{\mathcal{H}}(\tilde{\eta}_i^n) + \tilde{\mathbf{b}}_i^{x,n}. \quad (4.56)$$

It can be rewritten in the more compact form

$$\tilde{\mathcal{H}}(\tilde{\eta}^{n+1}) + \mathbf{T} \tilde{\eta}^n + \mathbf{1} = \tilde{\mathbf{b}}^n \quad (4.57)$$

and is solved using the Newton method of Brugnano and Casulli [20, 21]. For this purpose, in the case when the unlimited DG method is used in the wet cells  $\mathbf{T}$  is not an  $M$ -matrix, which is a necessary condition for the convergence of the Newton method. Consequently, at the first MOOD step the Newton method is stopped after the first iteration.

Moreover, the detector is activated also in those cells where the water depth is bounded between 0 and  $H_{\text{Lim}}$  in order to enforce the stability of the detector. In addition, in order to avoid NaN values the velocity  $u$  is computed from  $U$  and  $H$  as

$$u = \frac{UH}{H^2 + \epsilon}, \quad (4.58)$$

where  $\epsilon$  is a small number typically chosen between  $10^{-10}$  and  $10^{-5}$  in order to avoid division by zero.

## 4.2 Sub-cell finite volume limiting of staggered semi-implicit DG schemes in 2D

The core of this section is the extension of the limiter to two space dimensions and its application to the shallow water equations. We briefly recall the unlimited semi-implicit DG method in 2D [54], we expose the 2D semi-implicit sub-cell finite volume method and finally we describe the combined method including the limiter. Also here, we refer to the second appendix (section A.2) for the details of the computation of the matrices and tensors that will be introduced.

### 4.2.1 Governing equations of the 2D model

We consider the two-dimensional shallow water equations in conservative variables without friction terms. The system is composed of the continuity equation and the momentum equations in  $x$  and  $y$  direction and reads

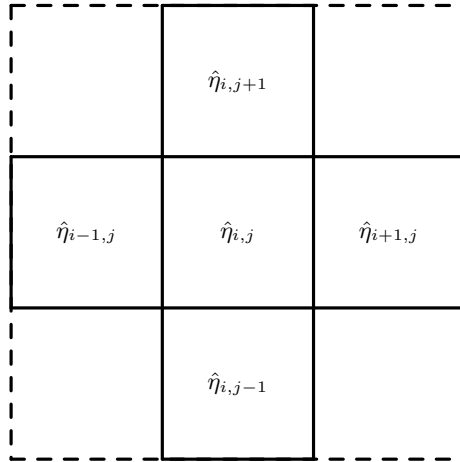
$$\begin{aligned} \frac{\partial \eta}{\partial t} + \frac{\partial U}{\partial x} + \frac{\partial V}{\partial y} &= 0, \\ \frac{\partial U}{\partial t} + \frac{\partial uU}{\partial x} + \frac{\partial vU}{\partial y} + gH \frac{\partial \eta}{\partial x} &= 0, \\ \frac{\partial V}{\partial t} + \frac{\partial uV}{\partial x} + \frac{\partial vV}{\partial y} + gH \frac{\partial \eta}{\partial y} &= 0, \end{aligned} \quad (4.59)$$



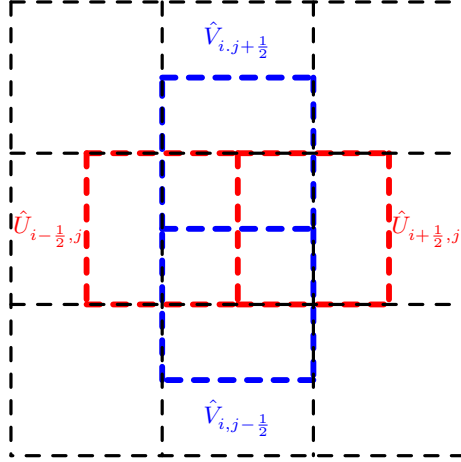
where the spatial coordinates are  $x$  and  $y$  and  $t$  is the time;  $\eta(x, y, t)$  is the elevation of the free surface,  $h(x, y)$  is the bottom which is assumed fixed in time,  $H(x, y, t) = \eta(x, y, t) + h(x, y)$  is the total water depth,  $u(x, y, t)$  and  $v(x, y, t)$  are the depth-averaged velocities respectively in  $x$  and  $y$  direction, while the depth-averaged momentum in  $x$  and  $y$  direction is defined as  $U = Hu$  and  $V = Hv$ . The computational domain is  $\Omega_{xy} = [x_L, x_R] \times [y_L, y_R]$ .

#### 4.2.2 Unlimited staggered semi-implicit DG method for the 2D shallow water equations

The main grid for the discretization of  $\Omega_{xy}$  is composed of  $N_x \times N_y$  elements and a generic control volume is denoted by  $T_{i,j} = [x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}] \times [y_{j-\frac{1}{2}}, y_{j+\frac{1}{2}}]$  (see Fig. 4.11). Likewise, two staggered grids are introduced, see Fig. 4.12, and their control volumes are denoted as  $T_{i+\frac{1}{2},j} = [x_i, x_{i+1}] \times [y_{j-\frac{1}{2}}, y_{j+\frac{1}{2}}]$  and  $T_{i,j+\frac{1}{2}} = [x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}] \times [y_j, y_{j+1}]$  for the  $x$  and  $y$  direction, respectively.



**Figure 4.11:** Computational grids for the 2D semi-implicit staggered DG scheme. Main grid for  $\eta$



**Figure 4.12:** Computational grids for the 2D semi-implicit staggered DG scheme. Staggered dual meshes for  $U$  and  $V$

In order to develop the semi-implicit DG scheme on staggered meshes we need the following sets of basis functions of degree  $P$

$$\begin{aligned}
 \phi(x, y) &= \varphi(\xi, \gamma) \quad \text{with} \quad x = x_i + \xi \Delta x \quad y = y_i + \gamma \Delta y \quad 0 \leq \xi, \gamma \leq 1 \\
 \psi(x, y) &= \varphi(\xi, \gamma) \quad \text{with} \quad x = x_{i+\frac{1}{2}} + \xi \Delta x \quad y = y_j + \gamma \Delta y \quad 0 \leq \xi, \gamma \leq 1. \\
 \omega(x, y) &= \varphi(\xi, \gamma) \quad \text{with} \quad x = x_i + \xi \Delta x \quad y = y_{j+\frac{1}{2}} + \gamma \Delta y \quad 0 \leq \xi, \gamma \leq 1.
 \end{aligned} \tag{4.60}$$

Similar to the 1D case, the free surface is located on the main grid and is represented within element  $T_{i,j}$  as follows

$$\eta_{i,j}(x, y, t^n) = \phi(\mathbf{x}, \mathbf{y}) \cdot \hat{\boldsymbol{\eta}}_{i,j}^n. \tag{4.61}$$

Then, the quantities that are defined on the staggered control volume  $T_{i+\frac{1}{2},j}$  are

$$\begin{aligned} U_{i+\frac{1}{2},j}(x, y, t^n) &= \psi(\mathbf{x}, \mathbf{y}) \cdot \hat{\mathbf{U}}_{i+\frac{1}{2},j}^n, \\ H_{i+\frac{1}{2},j}(x, y, t^n) &= \psi(\mathbf{x}, \mathbf{y}) \cdot \hat{\mathbf{H}}_{i+\frac{1}{2},j}^n, \\ u_{i+\frac{1}{2},j}(x, y, t^n) &= \psi(\mathbf{x}, \mathbf{y}) \cdot \hat{\mathbf{u}}_{i+\frac{1}{2},j}^n, \end{aligned} \quad (4.62)$$

while the quantities on the dual volumes  $T_{i,j+\frac{1}{2}}$  are

$$\begin{aligned} V_{i,j+\frac{1}{2}}(x, y, t^n) &= \omega(\mathbf{x}, \mathbf{y}) \cdot \hat{\mathbf{V}}_{i,j+\frac{1}{2}}^n, \\ H_{i,j+\frac{1}{2}}(x, y, t^n) &= \omega(\mathbf{x}, \mathbf{y}) \cdot \hat{\mathbf{H}}_{i,j+\frac{1}{2}}^n, \\ v_{i,j+\frac{1}{2}}(x, y, t^n) &= \omega(\mathbf{x}, \mathbf{y}) \cdot \hat{\mathbf{v}}_{i,j+\frac{1}{2}}^n. \end{aligned} \quad (4.63)$$

First we multiply the continuity equation by the vector of test functions  $\phi$  and integrate over a control volume  $T_{i,j}$  on the main grid

$$\int_{T_{i,j}} \phi \left( \frac{\partial \eta}{\partial t} + \frac{\partial U}{\partial x} + \frac{\partial V}{\partial y} \right) dx dy = 0. \quad (4.64)$$

After integration by parts we get the following semi-implicit discretization

$$\begin{aligned} \mathbf{M} \cdot (\hat{\boldsymbol{\eta}}_{i,j}^{n+1} - \hat{\boldsymbol{\eta}}_{i,j}^n) &+ \frac{\Delta t}{\Delta x} (\mathbf{R}_{\mathbf{j}}^{\mathbf{x},\text{DG}} \cdot \hat{\mathbf{U}}_{i+\frac{1}{2},j}^{n+\theta_{i+\frac{1}{2},j}} - \mathbf{L}_{\mathbf{j}}^{\mathbf{x},\text{DG}} \cdot \hat{\mathbf{U}}_{i-\frac{1}{2},j}^{n+\theta_{i-\frac{1}{2},j}}) \\ &+ \frac{\Delta t}{\Delta y} (\mathbf{R}_{\mathbf{j}}^{\mathbf{y},\text{DG}} \cdot \hat{\mathbf{V}}_{i,j+\frac{1}{2}}^{n+\theta_{i,j+\frac{1}{2}}} - \mathbf{L}_{\mathbf{j}}^{\mathbf{y},\text{DG}} \cdot \hat{\mathbf{V}}_{i,j-\frac{1}{2}}^{n+\theta_{i,j-\frac{1}{2}}}) = 0, \end{aligned} \quad (4.65)$$

where the  $\theta$ -method has been applied to spatial derivatives. Later on, also the momentum equations in  $x$  and  $y$  directions are multiplied by the test functions  $\psi$  and  $\omega$  and are integrated over the control volumes  $T_{i+\frac{1}{2},j}$  and  $T_{i,j+\frac{1}{2}}$ , respectively,

$$\int_{T_{i+\frac{1}{2},j}} \psi \left( \frac{\partial U}{\partial t} + \frac{\partial u U}{\partial x} + \frac{\partial v U}{\partial y} + g H \frac{\partial \eta}{\partial x} \right) dx dy = 0, \quad (4.66)$$

$$\int_{T_{i,j+\frac{1}{2}}} \omega \left( \frac{\partial V}{\partial t} + \frac{\partial u V}{\partial x} + \frac{\partial v V}{\partial y} + g H \frac{\partial \eta}{\partial y} \right) dx dy = 0, \quad (4.67)$$

obtaining the following formulation

$$\hat{\mathbf{U}}_{i+\frac{1}{2},j}^{n+1} = \widehat{\mathbf{GU}}_{i+\frac{1}{2},j}^n - g \frac{\Delta t}{\Delta x} \theta_{i+\frac{1}{2},j} \mathbf{M}^{-1} \cdot (\mathbf{R}_{\mathbf{u}}^{\mathbf{x},\text{DG}} \cdot \hat{\mathbf{H}}_{i+\frac{1}{2},j}^n \hat{\boldsymbol{\eta}}_{i+1,j}^{n+1} - \mathbf{L}_{\mathbf{u}}^{\mathbf{x},\text{DG}} \cdot \hat{\mathbf{H}}_{i+\frac{1}{2},j}^n \hat{\boldsymbol{\eta}}_{i,j}^{n+1}), \quad (4.68)$$

$$\hat{\mathbf{V}}_{i,j+\frac{1}{2}}^{n+1} = \widehat{\mathbf{G}\mathbf{V}}_{i,j+\frac{1}{2}}^n - g \frac{\Delta t}{\Delta y} \theta_{i,j+\frac{1}{2}} \mathbf{M}^{-1} \cdot (\mathbf{R}_{\mathbf{u}}^{\mathbf{y},\text{DG}} \cdot \hat{\mathbf{H}}_{i,j+\frac{1}{2}}^n \hat{\boldsymbol{\eta}}_{i,j+1}^{n+1} - \mathbf{L}_{\mathbf{u}}^{\mathbf{y},\text{DG}} \cdot \hat{\mathbf{H}}_{i,j+\frac{1}{2}}^n \hat{\boldsymbol{\eta}}_{i,j}^{n+1}), \quad (4.69)$$

where  $\widehat{\mathbf{G}\mathbf{U}}_{i+\frac{1}{2},j}^n$  and  $\widehat{\mathbf{G}\mathbf{V}}_{i,j+\frac{1}{2}}^n$  contain the nonlinear convective terms and the known terms at time  $t^n$

$$\widehat{\mathbf{G}\mathbf{U}}_{i+\frac{1}{2},j}^n = \widehat{\mathbf{F}\mathbf{U}}_{i+\frac{1}{2},j}^n - g \frac{\Delta t}{\Delta x} (1 - \theta_{i+\frac{1}{2},j}) \mathbf{M}^{-1} \cdot (\mathbf{R}_{\mathbf{u}}^{\mathbf{x},\text{DG}} \cdot \hat{\mathbf{H}}_{i+\frac{1}{2},j}^n \hat{\boldsymbol{\eta}}_{i+1,j}^n - \mathbf{L}_{\mathbf{u}}^{\mathbf{x},\text{DG}} \cdot \hat{\mathbf{H}}_{i+\frac{1}{2},j}^n \hat{\boldsymbol{\eta}}_{i,j}^n), \quad (4.70)$$

$$\widehat{\mathbf{G}\mathbf{V}}_{i,j+\frac{1}{2}}^n = \widehat{\mathbf{F}\mathbf{V}}_{i,j+\frac{1}{2}}^n - g \frac{\Delta t}{\Delta y} (1 - \theta_{i,j+\frac{1}{2}}) \mathbf{M}^{-1} \cdot (\mathbf{R}_{\mathbf{u}}^{\mathbf{y},\text{DG}} \cdot \hat{\mathbf{H}}_{i,j+\frac{1}{2}}^n \hat{\boldsymbol{\eta}}_{i,j+1}^n - \mathbf{L}_{\mathbf{u}}^{\mathbf{y},\text{DG}} \cdot \hat{\mathbf{H}}_{i,j+\frac{1}{2}}^n \hat{\boldsymbol{\eta}}_{i,j}^n). \quad (4.71)$$

Substituting  $\hat{\mathbf{V}}_{i+\frac{1}{2},j}^{n+1}$  and  $\hat{\mathbf{V}}_{i,j+\frac{1}{2}}^{n+1}$  into the discrete continuity equation (4.65) yields the block pentadiagonal system

$$\mathcal{L}_{i,j}^{\mathbf{x},\text{DG}} \cdot \hat{\boldsymbol{\eta}}_{i-1,j}^{n+1} + \mathcal{L}_{i,j}^{\mathbf{y},\text{DG}} \cdot \hat{\boldsymbol{\eta}}_{i,j-1}^{n+1} + \mathcal{C}_{i,j}^{\text{DG}} \cdot \hat{\boldsymbol{\eta}}_{i,j}^{n+1} + \mathcal{R}_{i,j}^{\mathbf{x},\text{DG}} \cdot \hat{\boldsymbol{\eta}}_{i+1,j}^{n+1} + \mathcal{R}_{i,j}^{\mathbf{y},\text{DG}} \cdot \hat{\boldsymbol{\eta}}_{i,j+1}^{n+1} = \hat{\mathbf{b}}_{i,j}^n, \quad (4.72)$$

which in practice we observe to be symmetric and positive-definite. Consequently, the unknown degrees of freedom of the free surface can be efficiently obtained via a matrix-free conjugate gradient method. Then, eq. (4.68) and eq. (4.69) provide the momentum updates, while the water depth  $H$  on the dual meshes is computed as

$$\hat{\mathbf{H}}_{i+\frac{1}{2},j} = \hat{\mathbf{h}}_{i+\frac{1}{2},j} + \mathbf{M}^{-1} \cdot (\mathbf{M}_{\mathbf{L}}^{\mathbf{x},\text{DG}} \cdot \hat{\boldsymbol{\eta}}_{i,j} + \mathbf{M}_{\mathbf{R}}^{\mathbf{x},\text{DG}} \cdot \hat{\boldsymbol{\eta}}_{i+1,j}), \quad (4.73)$$

$$\hat{\mathbf{H}}_{i,j+\frac{1}{2}} = \hat{\mathbf{h}}_{i,j+\frac{1}{2}} + \mathbf{M}^{-1} \cdot (\mathbf{M}_{\mathbf{L}}^{\mathbf{y},\text{DG}} \cdot \hat{\boldsymbol{\eta}}_{i,j} + \mathbf{M}_{\mathbf{R}}^{\mathbf{y},\text{DG}} \cdot \hat{\boldsymbol{\eta}}_{i,j+1}). \quad (4.74)$$

The nonlinear convective terms are discretized again on the main grid by an explicit two-dimensional RKDG scheme. Consequently, we define the following nonlinear operators used inside the Runge-Kutta time discretization

$$\mathbf{L}_{\mathbf{h}}^{\mathbf{x}}(\mathbf{u}, \mathbf{v}, \mathbf{U}, \mathbf{V})_{i,j} = -\frac{\Delta t}{\Delta x \Delta y} \mathbf{M}^{-1} \left( \int_{\partial T_{i,j}} \phi \mathbf{F}_{\mathbf{u}} \cdot \mathbf{n} dS - \int_{T_{i,j}} \nabla \phi \cdot \mathbf{F}_{\mathbf{u}} dx dy \right), \quad (4.75)$$

$$\mathbf{L}_{\mathbf{h}}^{\mathbf{y}}(\mathbf{u}, \mathbf{v}, \mathbf{U}, \mathbf{V})_{i,j} = -\frac{\Delta t}{\Delta x \Delta y} \mathbf{M}^{-1} \left( \int_{\partial T_{i,j}} \phi \mathbf{F}_{\mathbf{v}} \cdot \mathbf{n} dS - \int_{T_{i,j}} \nabla \phi \cdot \mathbf{F}_{\mathbf{v}} dx dy \right), \quad (4.76)$$

where  $\mathbf{F}_{\mathbf{u}} = [uU, vU]$  and  $\mathbf{F}_{\mathbf{v}} = [uV, vV]$  are the physical fluxes. In equations (4.75) and (4.76) and  $\hat{\mathbf{U}}_{i,j}^n$  and  $\hat{\mathbf{V}}_{i,j}^n$  have been computed on the main grid using

$$\hat{\mathbf{U}}_{i,j} = \mathbf{M}^{-1} \cdot (\mathbf{M}_{\mathbf{L}}^{\mathbf{x},\text{DG}} \cdot \hat{\mathbf{U}}_{i-\frac{1}{2},j} + \mathbf{M}_{\mathbf{R}}^{\mathbf{x},\text{DG}} \cdot \hat{\mathbf{U}}_{i+\frac{1}{2},j}), \quad (4.77)$$

$$\hat{\mathbf{V}}_{i,j} = \mathbf{M}^{-1} \cdot (\mathbf{M}_{\mathbf{L}}^{\mathbf{y},\text{DG}} \cdot \hat{\mathbf{U}}_{i,j-\frac{1}{2}} + \mathbf{M}_{\mathbf{R}}^{\mathbf{y},\text{DG}} \cdot \hat{\mathbf{V}}_{i,j+\frac{1}{2}}), \quad (4.78)$$

and  $\widehat{\mathbf{F}\mathbf{U}}_{i,j}^n$  and  $\widehat{\mathbf{F}\mathbf{V}}_{i,j}^n$  are projected back to the dual grid using the relations (4.73) and (4.74). The time step  $\Delta t$  is given by the CFL condition for explicit DG schemes based on the maximum absolute values of the convective speeds in the domain  $|u_{\max}|$  and  $|v_{\max}|$  with  $\text{CFL} < 1$ ,

$$\Delta t = \frac{\text{CFL}}{(2P+1)} \left( \frac{2|u_{\max}|}{\Delta x} + \frac{2|v_{\max}|}{\Delta y} \right)^{-1}. \quad (4.79)$$

### 4.2.3 Sub-cell formulation for the finite volume method for the 2D shallow water equations

Now we expose the semi-implicit sub-cell finite volume scheme for the 2D case. It is the natural two dimensional extension of the method proposed in sub-section 4.1.3. The computational grids and the control volumes  $T_{i,j}$ ,  $T_{i+\frac{1}{2},j}$  and  $T_{i,j+\frac{1}{2}}$  are the same introduced in the previous sub-section (see Figs. 4.13, 4.14, 4.15 and 4.16). For a given integer  $P$ , a sub-grid composed of  $(2P+1)^2$  finite volume sub-elements of size  $\Delta x_s = \Delta x/(2P+1)$  and  $\Delta y_s = \Delta y/(2P+1)$  is introduced (see Fig. 4.14). The degrees of freedom are allocated into a one dimensional vector of length  $(2P+1)^2$  ordered using the following relation  $s = i_1 + (i_2 - 1)(2P+1)$  with  $1 \leq i_1, i_2 \leq 2P+1$  and  $1 \leq s \leq (2P+1)^2$ , see again Fig. 4.14.

A semi-implicit discretization of the continuity equation on the main grid reads

$$\begin{aligned} \check{\eta}_{i,j}^{n+1} - \check{\eta}_{i,j}^n + \frac{\Delta t}{\Delta x} (\mathbf{R}^{\mathbf{x},\text{FV}} \cdot \check{\mathbf{U}}_{i+\frac{1}{2},j}^{n+\theta} - \mathbf{L}^{\mathbf{x},\text{FV}} \cdot \check{\mathbf{U}}_{i-\frac{1}{2},j}^{n+\theta}) \\ + \frac{\Delta t}{\Delta y} (\mathbf{R}^{\mathbf{y},\text{FV}} \cdot \check{\mathbf{V}}_{i,j+\frac{1}{2}}^{n+\theta} - \mathbf{L}^{\mathbf{y},\text{FV}} \cdot \check{\mathbf{V}}_{i,j-\frac{1}{2}}^{n+\theta}) = 0, \end{aligned} \quad (4.80)$$

and the discretized momentum equations on the dual meshes read

$$\check{\mathbf{U}}_{i+\frac{1}{2},j}^{n+1} = \check{\mathbf{G}}\mathbf{U}_{i+\frac{1}{2},j}^n - g \frac{\Delta t}{\Delta x} \theta_{i+\frac{1}{2},j} \check{\mathbf{H}}_{i+\frac{1}{2},j}^n \cdot (\mathbf{R}^{\mathbf{x},\text{FV}} \cdot \check{\eta}_{i+1,j}^{n+1} - \mathbf{L}^{\mathbf{x},\text{FV}} \cdot \check{\eta}_{i,j}^{n+1}), \quad (4.81)$$

$$\check{\mathbf{V}}_{i,j+\frac{1}{2}}^{n+1} = \check{\mathbf{G}}\mathbf{V}_{i,j+\frac{1}{2}}^n - g \frac{\Delta t}{\Delta y} \theta_{i,j+\frac{1}{2}} \check{\mathbf{H}}_{i,j+\frac{1}{2}}^n \cdot (\mathbf{R}^{\mathbf{y},\text{FV}} \cdot \check{\eta}_{i,j+1}^{n+1} - \mathbf{L}^{\mathbf{y},\text{FV}} \cdot \check{\eta}_{i,j}^{n+1}), \quad (4.82)$$

where  $\check{\mathbf{G}}\mathbf{U}_{i+\frac{1}{2},j}^n$  and  $\check{\mathbf{G}}\mathbf{V}_{i,j+\frac{1}{2}}^n$  as usual contain the explicit terms

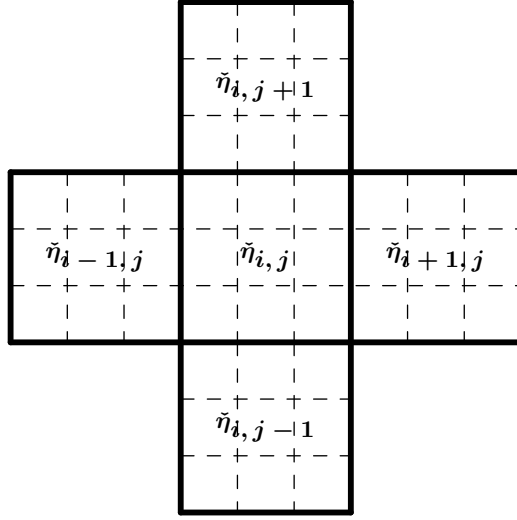
$$\check{\mathbf{G}}\mathbf{U}_{i+\frac{1}{2},j}^n = \check{\mathbf{F}}\mathbf{U}_{i+\frac{1}{2},j}^n - g \frac{\Delta t}{\Delta x} (1 - \theta_{i+\frac{1}{2},j}) \check{\mathbf{H}}_{i+\frac{1}{2},j}^n \cdot (\mathbf{R}^{\mathbf{x},\text{FV}} \cdot \check{\eta}_{i+1,j}^n - \mathbf{L}^{\mathbf{x},\text{FV}} \cdot \check{\eta}_{i,j}^n), \quad (4.83)$$

$$\check{\mathbf{G}}\mathbf{V}_{i,j+\frac{1}{2}}^n = \check{\mathbf{F}}\mathbf{V}_{i,j+\frac{1}{2}}^n - g \frac{\Delta t}{\Delta y} (1 - \theta_{i,j+\frac{1}{2}}) \check{\mathbf{H}}_{i,j+\frac{1}{2}}^n \cdot (\mathbf{R}^{\mathbf{y},\text{FV}} \cdot \check{\eta}_{i,j+1}^n - \mathbf{L}^{\mathbf{y},\text{FV}} \cdot \check{\eta}_{i,j}^n). \quad (4.84)$$

The block penta-diagonal system for the two-dimensional sub-cell finite volume reads

$$\mathcal{L}_{i,j}^{\mathbf{x},\text{FV}} \cdot \check{\eta}_{i-1,j}^{n+1} + \mathcal{L}_{i,j}^{\mathbf{y},\text{FV}} \cdot \check{\eta}_{i,j-1}^{n+1} + \mathcal{C}_{i,j}^{\text{FV}} \cdot \check{\eta}_{i,j}^{n+1} + \mathcal{R}_{i,j}^{\mathbf{x},\text{FV}} \cdot \check{\eta}_{i+1,j}^{n+1} + \mathcal{R}_{i,j}^{\mathbf{y},\text{FV}} \cdot \check{\eta}_{i,j+1}^{n+1} = \check{\mathbf{b}}_{i,j}^n, \quad (4.85)$$

obtained again after inserting the discrete momentum equations (4.81) and (4.82) into the continuity equation (4.80). The sub-cell averages of the water depth  $H$  inside the cells  $T_{i+\frac{1}{2},j}$  and  $T_{i,j+\frac{1}{2}}$



**Figure 4.13:** Main grid and stencil for the 2D semi-implicit sub-cell finite volume scheme.

are computed as follows

$$\check{H}_{i+\frac{1}{2},j} = \check{h}_{i+\frac{1}{2},j} + (M_L^{x,FV} \cdot \check{\eta}_{i,j} + M_R^{x,FV} \cdot \check{\eta}_{i+1,j}), \quad (4.86)$$

$$\check{H}_{i,j+\frac{1}{2}} = \check{h}_{i,j+\frac{1}{2}} + (M_L^{y,FV} \cdot \check{\eta}_{i,j} + M_R^{y,FV} \cdot \check{\eta}_{i,j+1}). \quad (4.87)$$

After projecting  $\check{U}_{i+\frac{1}{2},j}$  and  $\check{V}_{i,j+\frac{1}{2}}$  on the main grid via

$$\check{U}_{i,j} = (M_L^{x,FV} \cdot \check{U}_{i-\frac{1}{2},j} + M_R^{x,FV} \cdot \check{U}_{i+\frac{1}{2},j}), \quad (4.88)$$

$$\check{V}_{i,j} = (M_L^{y,FV} \cdot \check{V}_{i,j-\frac{1}{2}} + M_R^{y,FV} \cdot \check{V}_{i,j+\frac{1}{2}}), \quad (4.89)$$

it is possible to update the non linear convective terms  $\check{U}_{i,j}$  and  $\check{V}_{i,j}$  using a standard explicit finite volume method and a Rusanov flux at the cell interfaces. For this scheme, the CFL condition yields

$\check{\eta}_{i,j,7}$ $i_1 = 1, i_2 = 3$	$\check{\eta}_{i,j,8}$ $i_1 = 2, i_2 = 3$	$\check{\eta}_{i,j,9}$ $i_1 = 3, i_2 = 3$
$\check{\eta}_{i,j,4}$ $i_1 = 1, i_2 = 2$	$\check{\eta}_{i,j,5}$ $i_1 = 2, i_2 = 2$	$\check{\eta}_{i,j,6}$ $i_1 = 3, i_2 = 2$
$\check{\eta}_{i,j,1}$ $i_1 = 1, i_2 = 1$	$\check{\eta}_{i,j,2}$ $i_1 = 2, i_2 = 1$	$\check{\eta}_{i,j,3}$ $i_1 = 3, i_2 = 1$

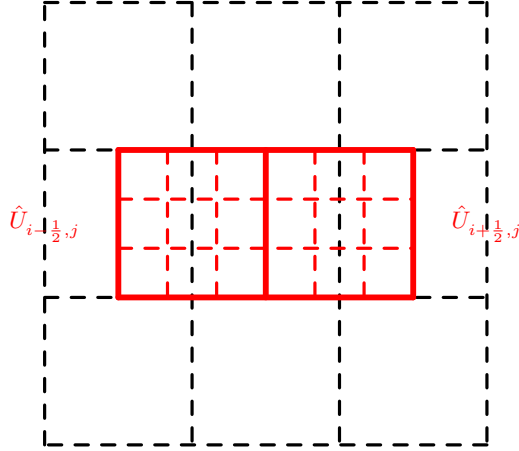
**Figure 4.14:** Ordering of the subcell degrees of freedom within one cell for the 2D semi-implicit sub-cell finite volume scheme

the following expression for  $\Delta t$

$$\Delta t = \text{CFL} \left( \frac{2|u_{\max}|}{\Delta x_s} + \frac{2|v_{\max}|}{\Delta y_s} \right)^{-1}. \quad (4.90)$$

#### 4.2.4 MOOD algorithm and detection criteria - 2D case

The extension to the 2D case of the the considerations made in Section 4.1.4 is straightforward, i.e. the MOOD paradigm is based on the same approach of Section 4.1.4. Also in the 2D case we need to compute properly the projection and reconstruction matrices  $\mathcal{P}^{xy}$  and  $\mathcal{W}^{xy}$  so that for a generic quantity  $q$  we have  $\check{\mathbf{q}}_{i,j}^n = \mathcal{P}^{xy} \cdot \hat{\mathbf{q}}_{i,j}^n$ ,  $\hat{\mathbf{q}}_{i,j}^n = \mathcal{W}^{xy} \cdot \check{\mathbf{q}}_{i,j}^n$ . In addition, we introduce also the operators  $\mathcal{W}_L^x$ ,  $\mathcal{W}_R^x$ ,  $\mathcal{W}_L^y$  and  $\mathcal{W}_R^y$  that reconstruct the DG polynomial from the piecewise constant data in the subdomains  $\Omega_L^x$ ,  $\Omega_R^x$ ,  $\Omega_L^y$  and  $\Omega_R^y$  of a control volume  $T_{i,j}$ , see Figures 4.17 and 4.18. We use the same detection criteria as in the 1D case, i.e. the positivity of the total water depth as



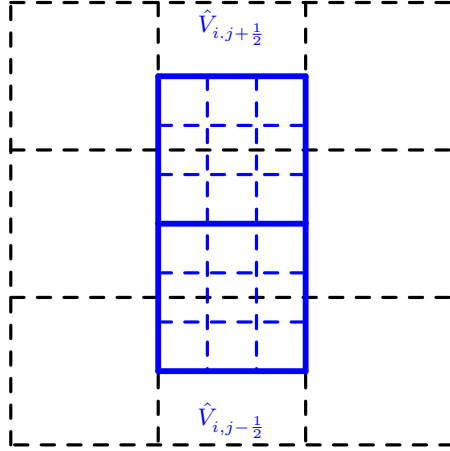
**Figure 4.15:** Staggered mesh in  $x$  direction for the 2D semi-implicit sub-cell finite volume scheme

well as the relaxed discrete maximum principle:

$$\hat{H}_{i,j,m}^{*,n+1} \geq 0, \quad \text{and} \quad \min_{\forall T_{k,l} \in \mathcal{V}_{i,j}} (\hat{q}_{k,l,m}^n) - \delta \leq \hat{q}_{i,j,m}^{*,n+1} \leq \max_{\forall T_{k,l} \in \mathcal{V}_{i,j}} (\hat{q}_{k,l,m}^n) + \delta, \quad (4.91)$$

with  $\mathcal{V}_{i,j} = \{T_{i-1,j-1}, T_{i,j-1}, T_{i+1,j-1}, T_{i-1,j}, T_{i,j}, T_{i+1,j}, T_{i-1,j+1}, T_{i,j+1}, T_{i+1,j+1}\}$  the set of Voronoi neighbours of  $T_{i,j}$ . The relaxation parameter  $\delta$  is computed in analogy to the 1D case. As detection quantities  $q$  in the DMP we use  $\eta$  as well as  $U$  and  $V$ . If a cell does not satisfy all the admissibility criteria in eq. (4.91), it is flagged as troubled cell by assigning a limiter status  $\beta_{i,j} = 1$ , while for all cells which do not need the limiting we set  $\beta_{i,j} = 0$ . Next, we need to limit properly also the corresponding cells of the staggered dual grids. For example, if the cell  $T_{i,j}$  is marked as troubled zone with  $\beta_{i,j} = 1$ , then also the four overlapping dual cells  $T_{i+\frac{1}{2},j}$ ,  $T_{i-\frac{1}{2},j}$ ,  $T_{i,j+\frac{1}{2}}$  and  $T_{i,j-\frac{1}{2}}$  need to be limited by setting  $\beta_{i+\frac{1}{2},j} = 1$ ,  $\beta_{i-\frac{1}{2},j} = 1$ ,  $\beta_{i,j+\frac{1}{2}} = 1$  and  $\beta_{i,j-\frac{1}{2}} = 1$ . Furthermore, according to the methodology adopted in the one dimensional case in section 4.1.4, in the first MOOD step the implicitness parameter  $\theta$  can be chosen between 0.5 and 1. Typically we chose  $\theta_{DG} = 0.55$  in the entire domain in order to reduce the numerical viscosity introduced by the time integration. However, at the second MOOD stage  $\theta$  is modified and imposed





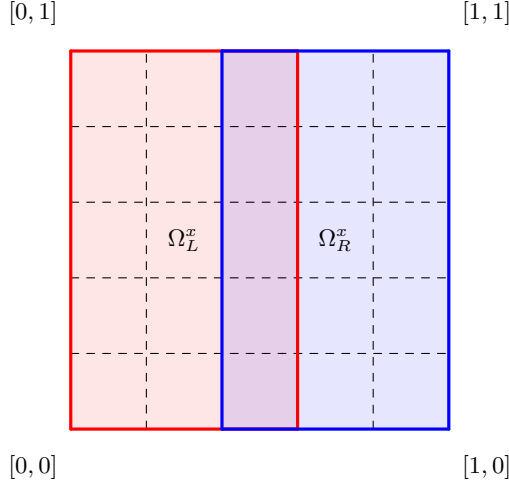
**Figure 4.16:** Staggered mesh in  $y$  direction for the 2D semi-implicit sub-cell finite volume scheme

equal to  $\theta = 1$  in the troubled cells, in order to avoid the generation of non-physical oscillations, while it is kept at  $\theta = \theta_{DG}$  in all the unlimited cells.

#### 4.2.5 Sub-cell limiting of the semi-implicit DG scheme for the 2D shallow water equations

For the continuity equation discretized in the limited case we consider the case where  $T_{i,j}^n$  is unlimited ( $\beta_{i,j} = 0$ ), but one of its edge neighbours is limited. We write the discrete continuity equation as follows

$$\mathbf{M} \cdot \left( \frac{\hat{\eta}_{i,j}^{n+1} - \hat{\eta}_i^n}{\Delta t} \right) + \frac{\Delta \hat{U}_{i,j}^{n+\theta_{i,j}}}{\Delta x} + \frac{\Delta \hat{V}_{i,j}^{n+\theta_{i,j}}}{\Delta y} = 0. \quad (4.92)$$



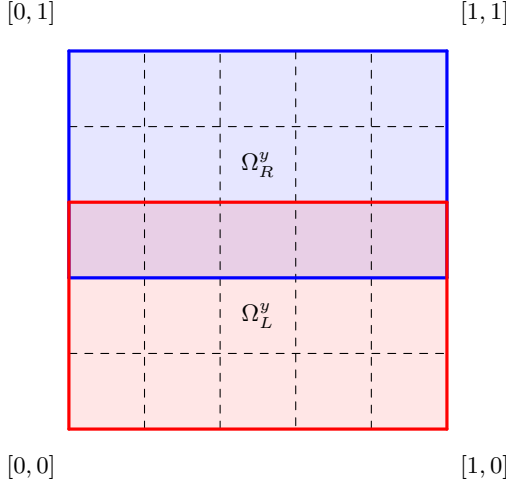
**Figure 4.17:** Subcell finite volumes in two space dimensions. Reference domains  $\Omega_L^x$  and  $\Omega_R^x$  for the reconstructions

We distinguish two cases for  $\Delta \hat{U}_{i,j}^{n+\theta_{i,j}}$ :

$$\Delta \hat{U}_{i,j}^{n+\theta_{i,j}} = \begin{cases} \mathbf{R}_j^{\mathbf{x},\text{Lim}} \cdot \tilde{U}_{i+\frac{1}{2},j}^{n+\theta_{i+\frac{1}{2},j}} - \mathbf{L}_j^{\mathbf{x},\text{DG}} \cdot \tilde{U}_{i-\frac{1}{2},j}^{n+\theta_{i-\frac{1}{2},j}} & \text{if } \beta_{i+\frac{1}{2},j} = 1, \beta_{i-\frac{1}{2},j} = 0, \\ \mathbf{R}_j^{\mathbf{x},\text{DG}} \cdot \tilde{U}_{i+\frac{1}{2},j}^{n+\theta_{i+\frac{1}{2},j}} - \mathbf{L}_j^{\mathbf{x},\text{Lim}} \cdot \tilde{U}_{i-\frac{1}{2},j}^{n+\theta_{i-\frac{1}{2},j}} & \text{if } \beta_{i+\frac{1}{2},j} = 0, \beta_{i-\frac{1}{2},j} = 1. \end{cases} \quad (4.93)$$

The first case happens when the cell  $T_{i+1,j}$  is troubled and consequently  $\beta_{i+1,j} = \beta_{i+\frac{1}{2},j} = 1$ ; on the contrary the other situation is when the control volume  $T_{i-1,j}$  is a troubled cell. Here we have introduced the two new operators  $\mathbf{R}_j^{\mathbf{x},\text{Lim}} = \mathbf{R}_j^{\mathbf{x},\text{DG}} \cdot \mathcal{W}_L^x$  and  $\mathbf{L}_j^{\mathbf{x},\text{Lim}} = \mathbf{L}_j^{\mathbf{x},\text{DG}} \cdot \mathcal{W}_R^x$ . Similarly, there are two possible cases for  $\Delta \hat{V}_{i,j}^{n+\theta_{i,j}}$ :

$$\Delta \hat{V}_{i,j}^{n+\theta_{i,j}} = \begin{cases} \mathbf{R}_j^{\mathbf{y},\text{Lim}} \cdot \tilde{V}_{i,j+\frac{1}{2}}^{n+\theta_{i,j+\frac{1}{2}}} - \mathbf{L}_j^{\mathbf{y},\text{DG}} \cdot \tilde{V}_{i,j-\frac{1}{2}}^{n+\theta_{i,j-\frac{1}{2}}} & \text{if } \beta_{i,j+\frac{1}{2}} = 1, \beta_{i,j-\frac{1}{2}} = 0, \\ \mathbf{R}_j^{\mathbf{y},\text{DG}} \cdot \tilde{V}_{i,j+\frac{1}{2}}^{n+\theta_{i,j+\frac{1}{2}}} - \mathbf{L}_j^{\mathbf{y},\text{Lim}} \cdot \tilde{V}_{i,j-\frac{1}{2}}^{n+\theta_{i,j-\frac{1}{2}}} & \text{if } \beta_{i,j+\frac{1}{2}} = 0, \beta_{i,j-\frac{1}{2}} = 1. \end{cases} \quad (4.94)$$



**Figure 4.18:** Subcell finite volumes in two space dimensions. Reference domains  $\Omega_L^y$  and  $\Omega_R^y$  for the reconstructions

In equation (4.94) the matrices  $\mathbf{R}_j^{\mathbf{y},\text{Lim}} = \mathbf{R}_j^{\mathbf{y},\text{DG}} \cdot \mathcal{W}_L^y$  and  $\mathbf{L}_j^{\mathbf{y},\text{Lim}} = \mathbf{L}_j^{\mathbf{y},\text{DG}} \cdot \mathcal{W}_R^y$  are used in order to properly combine the data representation of the sub-cell finite volume scheme with the one of the semi-implicit DG method. If neither the cell  $T_{i,j}$  is limited nor its four edge neighbours then the continuity equation is given by eq. (4.65) of the pure DG scheme, while if  $T_{i,j}$  is limited we use the continuity equation (4.80) of the sub-cell finite volume scheme.

Similar to the one dimensional case, for the momentum equation in  $x$  direction for the limited case we consider when  $T_{i+\frac{1}{2},j}$  is limited ( $\beta_{i+\frac{1}{2},j} = 1$ ) and also one of the two cells  $T_{i+1,j}$  and  $T_{i,j}$  is limited while the other is a non-troubled DG cell. Consequently, if  $\beta_{i+\frac{1}{2},j} = \beta_{i,j} = 1$  and  $\beta_{i+1,j} = 0$ ,  $\tilde{U}_{i+\frac{1}{2},j}^{n+1}$  reads as follows

$$\tilde{U}_{i+\frac{1}{2},j}^{n+1} = \tilde{F}\tilde{U}_{i+\frac{1}{2},j}^n - g \frac{\Delta t}{\Delta x} \tilde{H}_{i+\frac{1}{2},j}^n \cdot (\mathbf{R}_{\mathbf{u}}^{\mathbf{x},\text{Lim}} \cdot \hat{\boldsymbol{\eta}}_{i+1,j}^{n+\theta_{i+\frac{1}{2},j}} - \mathbf{L}^{\mathbf{x},\text{FV}} \cdot \hat{\boldsymbol{\eta}}_{i,j}^{n+\theta_{i+\frac{1}{2},j}})$$

with  $\mathbf{R}_{\mathbf{u}}^{\mathbf{x},\text{Lim}} = \mathbf{R}^{\mathbf{x},\text{FV}} \cdot \mathcal{P}^{xy}$

(4.95)

and for  $\beta_{i+\frac{1}{2},j} = \beta_{i+1,j} = 1$  and  $\beta_{i,j} = 0$ ,  $\tilde{U}_{i+\frac{1}{2},j}^{n+1}$  is computed as

$$\begin{aligned} \tilde{U}_{i+\frac{1}{2},j}^{n+1} &= \tilde{F}V_{i+\frac{1}{2},j}^n - g \frac{\Delta t}{\Delta x} \tilde{H}_{i+\frac{1}{2},j}^n \cdot (\mathbf{R}^{\mathbf{x},\text{FV}} \cdot \tilde{\boldsymbol{\eta}}_{i+1,j}^{n+\theta_{i+\frac{1}{2},j}} - \mathbf{L}_{\mathbf{u}}^{\mathbf{x},\text{Lim}} \cdot \tilde{\boldsymbol{\eta}}_{i,j}^{n+\theta_{i+\frac{1}{2},j}}) \\ \text{with } \mathbf{L}_{\mathbf{u}}^{\mathbf{x},\text{Lim}} &= \mathbf{L}^{\mathbf{x},\text{FV}} \cdot \mathcal{P}^{xy}. \end{aligned} \quad (4.96)$$

For the momentum equation in  $y$  direction, we can essentially repeat the same procedure. When  $T_{i,j+\frac{1}{2}}$  and  $T_{i,j}$  are limited and  $T_{i,j+1}$  is not ( $\beta_{i,j+\frac{1}{2}} = \beta_{i,j} = 1$ ,  $\beta_{i,j+1} = 0$ ) we have

$$\begin{aligned} \tilde{V}_{i,j+\frac{1}{2}}^{n+1} &= \tilde{F}V_{i,j+\frac{1}{2}}^n - g \frac{\Delta t}{\Delta y} \tilde{H}_{i,j+\frac{1}{2}}^n \cdot (\mathbf{R}_{\mathbf{u}}^{\mathbf{y},\text{Lim}} \cdot \tilde{\boldsymbol{\eta}}_{i,j+1}^{n+\theta_{i,j+\frac{1}{2}}} - \mathbf{L}^{\mathbf{y},\text{FV}} \cdot \tilde{\boldsymbol{\eta}}_{i,j}^{n+\theta_{i,j+\frac{1}{2}}}) \\ \text{with } \mathbf{R}_{\mathbf{u}}^{\mathbf{y},\text{Lim}} &= \mathbf{R}^{\mathbf{y},\text{FV}} \cdot \mathcal{P}^{xy}, \end{aligned} \quad (4.97)$$

while for the other limited case  $\beta_{i,j+\frac{1}{2}} = \beta_{i,j+1} = 1$  and  $\beta_{i,j} = 0$  the momentum equation in  $y$ -direction reads

$$\begin{aligned} \tilde{V}_{i,j+\frac{1}{2}}^{n+1} &= \tilde{F}V_{i,j+\frac{1}{2}}^n - g \frac{\Delta t}{\Delta y} \tilde{H}_{i,j+\frac{1}{2}}^n \cdot (\mathbf{R}^{\mathbf{y},\text{FV}} \cdot \tilde{\boldsymbol{\eta}}_{i,j+1}^{n+\theta_{i,j+\frac{1}{2}}} - \mathbf{L}_{\mathbf{u}}^{\mathbf{y},\text{Lim}} \cdot \tilde{\boldsymbol{\eta}}_{i,j}^{n+\theta_{i,j+\frac{1}{2}}}) \\ \text{with } \mathbf{L}_{\mathbf{u}}^{\mathbf{y},\text{Lim}} &= \mathbf{L}^{\mathbf{y},\text{FV}} \cdot \mathcal{P}^{xy}. \end{aligned} \quad (4.98)$$

The linear system for the final two dimensional algorithm, which includes the pure unlimited DG case as well as the sub-cell finite volume limiter, is given by inserting the discrete momentum equations into the discrete continuity equations and formally reads

$$\mathcal{L}_{i,j}^{x,\text{Lim}} \cdot \tilde{\boldsymbol{\eta}}_{i-1,j}^{n+1} + \mathcal{L}_{i,j}^{y,\text{Lim}} \cdot \tilde{\boldsymbol{\eta}}_{i,j-1}^{n+1} + \mathcal{C}_{i,j}^{\text{Lim}} \cdot \tilde{\boldsymbol{\eta}}_{i,j}^{n+1} + \mathcal{R}_{i,j}^{x,\text{Lim}} \cdot \tilde{\boldsymbol{\eta}}_{i+1,j}^{n+1} + \mathcal{R}_{i,j}^{y,\text{Lim}} \cdot \tilde{\boldsymbol{\eta}}_{i,j+1}^{n+1} = \tilde{\mathbf{b}}_{i,j}^n, \quad (4.99)$$

where  $\tilde{\boldsymbol{\eta}}_{i,j}$  is interpreted as the set of the degrees of freedom of the DG method if  $\beta_{i,j} = 0$  or as the set of sub-cell averages inside the control volume  $T_{i,j}$  if  $\beta_{i,j} = 1$ , i.e.

$$\tilde{\boldsymbol{\eta}}_{i,j} = \begin{cases} \hat{\boldsymbol{\eta}}_{i,j} & \text{if } \beta_{i,j} = 0, \\ \tilde{\boldsymbol{\eta}}_{i,j} & \text{if } \beta_{i,j} = 1. \end{cases} \quad (4.100)$$

Furthermore, numerical experiments have shown that the system appears to be symmetric and positive definite, hence it is possible to obtain the solution of the free surface using a matrix-free conjugate gradient method. However, we stress that we do not have a rigorous mathematical proof of this property yet and further theoretical investigations are still necessary. For a better understanding, it easier to write  $\mathcal{C}_{i,j}^{\text{Lim}}$  as the sum of three contributions

$$\mathcal{C}_{i,j}^{\text{Lim}} = \mathcal{C}_{i,j}^{0,\text{Lim}} + \mathcal{C}_{i,j}^{x,\text{Lim}} + \mathcal{C}_{i,j}^{y,\text{Lim}}. \quad (4.101)$$

As in the 1D case,  $\mathcal{C}_{i,j}^{0,\text{Lim}}$  is the element mass matrix if  $T_{i,j}$  is unlimited or the identity matrix if  $T_{i,j}$  is limited, i.e.

$$\mathcal{C}_{i,j}^{0,\text{Lim}} = \begin{cases} \mathbf{M} & \text{if } \beta_{i,j} = 0, \\ \mathbf{I} & \text{if } \beta_{i,j} = 1. \end{cases} \quad (4.102)$$

#### 4.2 Sub-cell finite volume limiting of staggered semi-implicit DG schemes in 2D

The other contributions  $C_{i,j}^{x,\text{Lim}}$  and  $C_{i,j}^{y,\text{Lim}}$  are computed considering four different combinations

$$C_{i,j}^{x,\text{Lim}} = g \frac{\Delta t^2}{\Delta x^2} \begin{cases} \theta_{i-\frac{1}{2},j}^2 \mathbf{L}_J^{\mathbf{x},\text{DG}} \cdot \mathbf{M}^{-1} \cdot \mathbf{R}_u^{\mathbf{x},\text{DG}} \cdot \hat{\mathbf{H}}_{i-\frac{1}{2},j}^n \\ + \theta_{i+\frac{1}{2},j}^2 \mathbf{R}_J^{\mathbf{x},\text{DG}} \cdot \mathbf{M}^{-1} \cdot \mathbf{L}_u^{\mathbf{x},\text{DG}} \cdot \hat{\mathbf{H}}_{i+\frac{1}{2},j}^n \\ \text{if } \beta_{i-\frac{1}{2},j} = \beta_{i+\frac{1}{2},j} = 0, \\ \theta_{i-\frac{1}{2},j}^2 \mathbf{L}_J^{\mathbf{x},\text{FV}} \cdot \tilde{\mathcal{H}}_{i-\frac{1}{2},j}^n \cdot \mathbf{R}_u^{\mathbf{x},\text{FV}} \\ + \theta_{i+\frac{1}{2},j}^2 \mathbf{R}_J^{\mathbf{x},\text{FV}} \cdot \tilde{\mathcal{H}}_{i+\frac{1}{2},j}^n \cdot \mathbf{L}_u^{\mathbf{x},\text{FV}} \\ \text{if } \beta_{i-\frac{1}{2},j} = \beta_{i+\frac{1}{2},j} = 1, \\ \theta_{i-\frac{1}{2},j}^2 \mathbf{L}_J^{\mathbf{x},\text{DG}} \cdot \mathbf{M}^{-1} \cdot \mathbf{R}_u^{\mathbf{x},\text{DG}} \cdot \hat{\mathbf{H}}_{i-\frac{1}{2},j}^n \\ + \theta_{i+\frac{1}{2},j}^2 \mathbf{R}_J^{\mathbf{x},\text{Lim}} \cdot \tilde{\mathcal{H}}_{i+\frac{1}{2},j}^n \cdot \mathbf{L}_u^{\mathbf{x},\text{Lim}} \\ \text{if } \beta_{i-\frac{1}{2},j} = 0, \beta_{i+\frac{1}{2},j} = 1, \\ \theta_{i-\frac{1}{2},j}^2 \mathbf{L}_J^{\mathbf{x},\text{Lim}} \cdot \tilde{\mathcal{H}}_{i-\frac{1}{2},j}^n \cdot \mathbf{R}_u^{\mathbf{x},\text{Lim}} \\ + \theta_{i+\frac{1}{2},j}^2 \mathbf{R}_J^{\mathbf{x},\text{DG}} \cdot \mathbf{M}^{-1} \cdot \mathbf{L}_u^{\mathbf{x},\text{DG}} \cdot \hat{\mathbf{H}}_{i+\frac{1}{2},j}^n \\ \text{if } \beta_{i+\frac{1}{2},j} = 0, \beta_{i-\frac{1}{2},j} = 1, \end{cases} \quad (4.103)$$

$$C_{i,j}^{y,\text{Lim}} = g \frac{\Delta t^2}{\Delta y^2} \begin{cases} \theta_{i,j-\frac{1}{2}}^2 \mathbf{L}_J^{\mathbf{y},\text{DG}} \cdot \mathbf{M}^{-1} \mathbf{R}_u^{\mathbf{y},\text{DG}} \cdot \hat{\mathbf{H}}_{i,j-\frac{1}{2}}^n \\ + \theta_{i,j+\frac{1}{2}}^2 \mathbf{R}_J^{\mathbf{y},\text{DG}} \cdot \mathbf{M}^{-1} \cdot \mathbf{L}_u^{\mathbf{y},\text{DG}} \cdot \hat{\mathbf{H}}_{i,j+\frac{1}{2}}^n \\ \text{if } \beta_{i,j-\frac{1}{2}} = \beta_{i,j+\frac{1}{2}} = 0, \\ \theta_{i,j-\frac{1}{2}}^2 \mathbf{L}_J^{\mathbf{y},\text{FV}} \cdot \tilde{\mathcal{H}}_{i,j-\frac{1}{2}}^n \cdot \mathbf{R}_u^{\mathbf{y},\text{FV}} \\ + \theta_{i,j+\frac{1}{2}}^2 \mathbf{R}_J^{\mathbf{y},\text{FV}} \cdot \tilde{\mathcal{H}}_{i,j+\frac{1}{2}}^n \cdot \mathbf{L}_u^{\mathbf{y},\text{FV}} \\ \text{if } \beta_{i,j-\frac{1}{2}} = \beta_{i,j+\frac{1}{2}} = 1, \\ \theta_{i,j-\frac{1}{2}}^2 \mathbf{L}_J^{\mathbf{y},\text{DG}} \cdot \mathbf{M}^{-1} \cdot \mathbf{R}_u^{\mathbf{y},\text{DG}} \cdot \hat{\mathbf{H}}_{i,j-\frac{1}{2}}^n \\ + \theta_{i,j+\frac{1}{2}}^2 \mathbf{R}_J^{\mathbf{y},\text{Lim}} \cdot \tilde{\mathcal{H}}_{i,j+\frac{1}{2}}^n \cdot \mathbf{L}_u^{\mathbf{y},\text{Lim}} \\ \text{if } \beta_{i,j-\frac{1}{2}} = 0, \beta_{i,j+\frac{1}{2}} = 1, \\ \theta_{i,j-\frac{1}{2}}^2 \mathbf{L}_J^{\mathbf{y},\text{Lim}} \cdot \tilde{\mathcal{H}}_{i,j-\frac{1}{2}}^n \cdot \mathbf{R}_u^{\mathbf{y},\text{Lim}} \\ + \theta_{i,j+\frac{1}{2}}^2 \mathbf{R}_J^{\mathbf{y},\text{DG}} \cdot \mathbf{M}^{-1} \cdot \mathbf{L}_u^{\mathbf{y},\text{DG}} \cdot \hat{\mathbf{H}}_{i,j+\frac{1}{2}}^n \\ \text{if } \beta_{i,j+\frac{1}{2}} = 0, \beta_{i,j-\frac{1}{2}} = 1. \end{cases} \quad (4.104)$$

The expressions for the blocks  $\mathcal{L}_{i,j}^{x,\text{Lim}}$ ,  $\mathcal{L}_{i,j}^{y,\text{Lim}}$ ,  $\mathcal{R}_{i,j}^{x,\text{Lim}}$  and  $\mathcal{R}_{i,j}^{y,\text{Lim}}$  are

$$\mathcal{L}_{i,j}^{x,\text{Lim}} = -g \frac{\Delta t^2}{\Delta x^2} \theta_{i-\frac{1}{2},j}^2 \begin{cases} \mathbf{L}_j^{\mathbf{x},\text{DG}} \cdot \mathbf{M}^{-1} \cdot \mathbf{L}_u^{\mathbf{x},\text{DG}} \cdot \hat{\mathbf{H}}_{i-\frac{1}{2},j}^n & \text{if } \beta_{i-1,j} = \beta_{i,j} = 0, \\ \mathbf{L}_j^{\mathbf{x},\text{FV}} \cdot \tilde{\mathcal{H}}_{i-\frac{1}{2},j}^n \cdot \mathbf{L}_j^{\mathbf{x},\text{FV}} & \text{if } \beta_{i-1,j} = \beta_{i,j} = 1, \\ \mathbf{L}_j^{\mathbf{x},\text{FV}} \cdot \tilde{\mathcal{H}}_{i-\frac{1}{2},j}^n \cdot \mathbf{L}_u^{\mathbf{x},\text{Lim}} & \text{if } \beta_{i-1,j} = 1, \beta_{i,j} = 0, \\ \mathbf{L}_j^{\mathbf{x},\text{Lim}} \cdot \tilde{\mathcal{H}}_{i-\frac{1}{2},j}^n \cdot \mathbf{L}_j^{\mathbf{x},\text{FV}} & \text{if } \beta_{i-1,j} = 0, \beta_{i,j} = 1, \end{cases} \quad (4.105)$$

$$\mathcal{L}_{i,j}^{y,\text{Lim}} = -g \frac{\Delta t^2}{\Delta y^2} \theta_{i,j-\frac{1}{2}}^2 \begin{cases} \mathbf{L}_j^{\mathbf{y},\text{DG}} \cdot \mathbf{M}^{-1} \cdot \mathbf{L}_u^{\mathbf{y},\text{DG}} \cdot \hat{\mathbf{H}}_{i,j-\frac{1}{2}}^n & \text{if } \beta_{i,j-1} = \beta_{i,j} = 0, \\ \mathbf{L}_j^{\mathbf{y},\text{FV}} \cdot \tilde{\mathcal{H}}_{i,j-\frac{1}{2}}^n \cdot \mathbf{L}_j^{\mathbf{y},\text{FV}} & \text{if } \beta_{i,j-1} = \beta_{i,j} = 1, \\ \mathbf{L}_j^{\mathbf{y},\text{FV}} \cdot \tilde{\mathcal{H}}_{i,j-\frac{1}{2}}^n \cdot \mathbf{L}_u^{\mathbf{y},\text{Lim}} & \text{if } \beta_{i,j-1} = 1, \beta_{i,j} = 0, \\ \mathbf{L}_j^{\mathbf{y},\text{Lim}} \cdot \tilde{\mathcal{H}}_{i,j-\frac{1}{2}}^n \cdot \mathbf{L}_j^{\mathbf{y},\text{FV}} & \text{if } \beta_{i,j-1} = 0, \beta_{i,j} = 1, \end{cases} \quad (4.106)$$

$$\mathcal{R}_{i,j}^{x,\text{Lim}} = -g \frac{\Delta t^2}{\Delta x^2} \theta_{i+\frac{1}{2},j}^2 \begin{cases} \mathbf{R}_j^{\mathbf{x},\text{DG}} \cdot \mathbf{M}^{-1} \cdot \mathbf{R}_u^{\mathbf{x},\text{DG}} \cdot \hat{\mathbf{H}}_{i+\frac{1}{2},j}^n & \text{if } \beta_{i+1,j} = \beta_{i,j} = 0, \\ \mathbf{R}_j^{\mathbf{x},\text{FV}} \cdot \tilde{\mathcal{H}}_{i+\frac{1}{2},j}^n \cdot \mathbf{R}_j^{\mathbf{x},\text{FV}} & \text{if } \beta_{i+1,j} = \beta_{i,j} = 1, \\ \mathbf{R}_j^{\mathbf{x},\text{FV}} \cdot \tilde{\mathcal{H}}_{i+\frac{1}{2},j}^n \cdot \mathbf{R}_u^{\mathbf{x},\text{Lim}} & \text{if } \beta_{i+1,j} = 1, \beta_{i,j} = 0, \\ \mathbf{R}_j^{\mathbf{x},\text{Lim}} \cdot \tilde{\mathcal{H}}_{i+\frac{1}{2},j}^n \cdot \mathbf{R}_j^{\mathbf{x},\text{FV}} & \text{if } \beta_{i+1,j} = 0, \beta_{i,j} = 1, \end{cases} \quad (4.107)$$

$$\mathcal{R}_{i,j}^{y,\text{Lim}} = -g \frac{\Delta t^2}{\Delta y^2} \theta_{i,j+\frac{1}{2}}^2 \begin{cases} \mathbf{R}_j^{\mathbf{y},\text{DG}} \cdot \mathbf{M}^{-1} \cdot \mathbf{R}_u^{\mathbf{y},\text{DG}} \cdot \hat{\mathbf{H}}_{i,j+\frac{1}{2}}^n & \text{if } \beta_{i,j+1} = \beta_{i,j} = 0, \\ \mathbf{R}_j^{\mathbf{y},\text{FV}} \cdot \tilde{\mathcal{H}}_{i,j+\frac{1}{2}}^n \cdot \mathbf{R}_j^{\mathbf{y},\text{FV}} & \text{if } \beta_{i,j+1} = \beta_{i,j} = 1, \\ \mathbf{R}_j^{\mathbf{y},\text{FV}} \cdot \tilde{\mathcal{H}}_{i,j+\frac{1}{2}}^n \cdot \mathbf{R}_u^{\mathbf{y},\text{Lim}} & \text{if } \beta_{i,j+1} = 1, \beta_{i,j} = 0, \\ \mathbf{R}_j^{\mathbf{y},\text{Lim}} \cdot \tilde{\mathcal{H}}_{i,j+\frac{1}{2}}^n \cdot \mathbf{R}_j^{\mathbf{y},\text{FV}} & \text{if } \beta_{i,j+1} = 0, \beta_{i,j} = 1. \end{cases} \quad (4.108)$$

Also the right hand side terms is split in three contributions for simplicity

$$\tilde{\mathbf{b}}_{i,j}^n = \tilde{\mathbf{b}}_{i,j}^{0,n} - \tilde{\mathbf{b}}_{i,j}^{x,n} - \tilde{\mathbf{b}}_{i,j}^{y,n} \quad (4.109)$$

where

$$\tilde{\mathbf{b}}_{i,j}^{0,n} = \begin{cases} \mathbf{M} \cdot \hat{\boldsymbol{\eta}}_{i,j}^n & \text{if } \beta_{i,j} = 0, \\ \tilde{\boldsymbol{\eta}}_{i,j}^n & \text{if } \beta_{i,j} = 1, \end{cases} \quad (4.110)$$

$$\tilde{b}_{i,j}^{x,n} = \frac{\Delta t}{\Delta x} \left\{ \begin{array}{l} (\theta_{i+\frac{1}{2},j} \mathbf{R}_j^{\mathbf{x},\text{DG}} \cdot \widehat{\mathbf{GU}}_{i+\frac{1}{2},j}^n - \theta_{i-\frac{1}{2},j} \mathbf{L}_j^{\mathbf{x},\text{DG}} \cdot \widehat{\mathbf{GU}}_{i-\frac{1}{2},j}^n) + \\ ((1 - \theta_{i+\frac{1}{2},j}) \mathbf{R}_j^{\mathbf{x},\text{DG}} \cdot \check{\mathbf{U}}_{i+\frac{1}{2},j}^n - (1 - \theta_{i-\frac{1}{2},j}) \mathbf{L}_j^{\mathbf{x},\text{DG}} \cdot \check{\mathbf{U}}_{i-\frac{1}{2},j}^n) \\ \text{if } \beta_{i,j} = \beta_{i-\frac{1}{2},j} = \beta_{i+\frac{1}{2},j} = 0, \\ (\theta_{i+\frac{1}{2},j} \mathbf{R}_j^{\mathbf{x},\text{FV}} \cdot \check{\mathbf{GU}}_{i+\frac{1}{2},j}^n - \theta_{i-\frac{1}{2},j} \mathbf{L}_j^{\mathbf{x},\text{FV}} \cdot \check{\mathbf{GU}}_{i-\frac{1}{2},j}^n) + \\ ((1 - \theta_{i+\frac{1}{2},j}) \mathbf{R}_j^{\mathbf{x},\text{FV}} \cdot \check{\mathbf{U}}_{i+\frac{1}{2},j}^n - (1 - \theta_{i-\frac{1}{2},j}) \mathbf{L}_j^{\mathbf{x},\text{FV}} \cdot \check{\mathbf{U}}_{i-\frac{1}{2},j}^n) \\ \text{if } \beta_{i,j} = \beta_{i-\frac{1}{2},j} = \beta_{i+\frac{1}{2},j} = 1, \\ (\theta_{i+\frac{1}{2},j} \mathbf{R}_j^{\mathbf{x},\text{Lim}} \cdot \check{\mathbf{GU}}_{i+\frac{1}{2},j}^n - \theta_{i-\frac{1}{2},j} \mathbf{L}_j^{\mathbf{x},\text{DG}} \cdot \widehat{\mathbf{GU}}_{i-\frac{1}{2},j}^n) + \\ ((1 - \theta_{i+\frac{1}{2},j}) \mathbf{R}_j^{\mathbf{x},\text{Lim}} \cdot \check{\mathbf{U}}_{i+\frac{1}{2},j}^n - (1 - \theta_{i-\frac{1}{2},j}) \mathbf{L}_j^{\mathbf{x},\text{DG}} \cdot \check{\mathbf{U}}_{i-\frac{1}{2},j}^n) \\ \text{if } \beta_{i,j} = \beta_{i-\frac{1}{2},j} = 0, \beta_{i+\frac{1}{2},j} = 1, \\ (\theta_{i+\frac{1}{2},j} \mathbf{R}_j^{\mathbf{x},\text{DG}} \cdot \widehat{\mathbf{GU}}_{i+\frac{1}{2},j}^n - \theta_{i-\frac{1}{2},j} \mathbf{L}_j^{\mathbf{x},\text{Lim}} \cdot \check{\mathbf{GU}}_{i-\frac{1}{2},j}^n) + \\ ((1 - \theta_{i+\frac{1}{2},j}) \mathbf{R}_j^{\mathbf{x},\text{DG}} \cdot \check{\mathbf{U}}_{i+\frac{1}{2},j}^n - (1 - \theta_{i-\frac{1}{2},j}) \mathbf{L}_j^{\mathbf{x},\text{Lim}} \cdot \check{\mathbf{U}}_{i-\frac{1}{2},j}^n) \\ \text{if } \beta_{i,j} = \beta_{i+\frac{1}{2},j} = 0, \beta_{i-\frac{1}{2},j} = 1, \end{array} \right. \quad (4.111)$$

$$\tilde{b}_{i,j}^{y,n} = \frac{\Delta t}{\Delta y} \left\{ \begin{array}{l} (\theta_{i,j+\frac{1}{2}} \mathbf{R}_j^{\mathbf{y},\text{DG}} \cdot \widehat{\mathbf{GV}}_{i,j+\frac{1}{2}}^n - \theta_{i,j-\frac{1}{2}} \mathbf{L}_j^{\mathbf{y},\text{DG}} \cdot \widehat{\mathbf{GV}}_{i,j-\frac{1}{2}}^n) + \\ + ((1 - \theta_{i,j+\frac{1}{2}}) \mathbf{R}_j^{\mathbf{y},\text{DG}} \cdot \check{\mathbf{V}}_{i,j+\frac{1}{2}}^n - (1 - \theta_{i,j-\frac{1}{2}}) \mathbf{L}_j^{\mathbf{y},\text{DG}} \cdot \check{\mathbf{V}}_{i,j-\frac{1}{2}}^n) \\ \text{if } \beta_{i,j} = \beta_{i,j-\frac{1}{2}} = \beta_{i,j+\frac{1}{2}} = 0, \\ (\theta_{i,j+\frac{1}{2}} \mathbf{R}_j^{\mathbf{y},\text{FV}} \cdot \check{\mathbf{GV}}_{i,j+\frac{1}{2}}^n - \theta_{i,j-\frac{1}{2}} \mathbf{L}_j^{\mathbf{y},\text{FV}} \cdot \check{\mathbf{GV}}_{i,j-\frac{1}{2}}^n) + \\ + ((1 - \theta_{i,j+\frac{1}{2}}) \mathbf{R}_j^{\mathbf{y},\text{FV}} \cdot \check{\mathbf{V}}_{i,j+\frac{1}{2}}^n - (1 - \theta_{i,j-\frac{1}{2}}) \mathbf{L}_j^{\mathbf{y},\text{FV}} \cdot \check{\mathbf{V}}_{i,j-\frac{1}{2}}^n) \\ \text{if } \beta_{i,j} = \beta_{i,j-\frac{1}{2}} = \beta_{i,j+\frac{1}{2}} = 1, \\ (\theta_{i,j+\frac{1}{2}} \mathbf{R}_j^{\mathbf{y},\text{Lim}} \cdot \check{\mathbf{GV}}_{i,j+\frac{1}{2}}^n - \theta_{i,j-\frac{1}{2}} \mathbf{L}_j^{\mathbf{y},\text{DG}} \cdot \widehat{\mathbf{GV}}_{i,j-\frac{1}{2}}^n) + \\ + ((1 - \theta_{i,j+\frac{1}{2}}) \mathbf{R}_j^{\mathbf{y},\text{Lim}} \cdot \check{\mathbf{V}}_{i,j+\frac{1}{2}}^n - (1 - \theta_{i,j-\frac{1}{2}}) \mathbf{L}_j^{\mathbf{y},\text{DG}} \cdot \check{\mathbf{V}}_{i,j-\frac{1}{2}}^n) \\ \text{if } \beta_{i,j} = \beta_{i,j-\frac{1}{2}} = 0, \beta_{i,j+\frac{1}{2}} = 1, \\ (\theta_{i,j+\frac{1}{2}} \mathbf{R}_j^{\mathbf{y},\text{DG}} \cdot \widehat{\mathbf{GV}}_{i,j+\frac{1}{2}}^n - \theta_{i,j-\frac{1}{2}} \mathbf{L}_j^{\mathbf{y},\text{Lim}} \cdot \check{\mathbf{GV}}_{i,j-\frac{1}{2}}^n) + \\ + ((1 - \theta_{i,j+\frac{1}{2}}) \mathbf{R}_j^{\mathbf{y},\text{DG}} \cdot \check{\mathbf{V}}_{i,j+\frac{1}{2}}^n - (1 - \theta_{i,j-\frac{1}{2}}) \mathbf{L}_j^{\mathbf{y},\text{Lim}} \cdot \check{\mathbf{V}}_{i,j-\frac{1}{2}}^n) \\ \text{if } \beta_{i,j} = \beta_{i,j+\frac{1}{2}} = 0, \beta_{i,j-\frac{1}{2}} = 1. \end{array} \right. \quad (4.112)$$

Once the new free surface elevation is known from the solution of (4.99), the momentum can be updated and the new water depth  $H$  needs to be computed on the dual meshes via

$$\tilde{H}_{i+\frac{1}{2},j} = \begin{cases} \tilde{h}_{i+\frac{1}{2},j} + \mathbf{M}^{-1} \cdot (\mathbf{M}_{\mathbf{L}}^{\mathbf{x},\text{DG}} \cdot \tilde{\eta}_{i,j} + \mathbf{M}_{\mathbf{R}}^{\mathbf{x},\text{DG}} \cdot \tilde{\eta}_{i+1,j}) & \text{if } \beta_{i+\frac{1}{2},j} = \beta_{i,j} = \beta_{i+1,j} = 0, \\ \tilde{h}_{i+\frac{1}{2},j} + (\mathbf{M}_{\mathbf{L}}^{\mathbf{x},\text{FV}} \cdot \tilde{\eta}_{i,j} + \mathbf{M}_{\mathbf{R}}^{\mathbf{x},\text{FV}} \cdot \tilde{\eta}_{i+1,j}) & \text{if } \beta_{i+\frac{1}{2},j} = \beta_{i,j} = \beta_{i+1,j} = 1, \\ \tilde{h}_{i+\frac{1}{2},j} + (\mathbf{M}_{\mathbf{L}}^{\mathbf{x},\text{FV}} \cdot \mathcal{P}^{xy} \cdot \tilde{\eta}_{i,j} + \mathbf{M}_{\mathbf{R}}^{\mathbf{x},\text{FV}} \cdot \tilde{\eta}_{i+1,j}) & \text{if } \beta_{i+\frac{1}{2},j} = \beta_{i+1,j} = 1, \beta_{i,j} = 0, \\ \tilde{h}_{i+\frac{1}{2},j} + (\mathbf{M}_{\mathbf{L}}^{\mathbf{x},\text{FV}} \cdot \tilde{\eta}_{i,j} + \mathbf{M}_{\mathbf{R}}^{\mathbf{x},\text{FV}} \cdot \mathcal{P}^{xy} \cdot \tilde{\eta}_{i+1,j}) & \text{if } \beta_{i+\frac{1}{2},j} = \beta_{i,j} = 1, \beta_{i+1,j} = 0, \end{cases} \quad (4.113)$$

$$\tilde{H}_{i,j+\frac{1}{2}} = \begin{cases} \tilde{h}_{i,j+\frac{1}{2}} + \mathbf{M}^{-1} \cdot (\mathbf{M}_{\mathbf{L}}^{\mathbf{y},\text{DG}} \cdot \tilde{\eta}_{i,j} + \mathbf{M}_{\mathbf{R}}^{\mathbf{y},\text{DG}} \cdot \tilde{\eta}_{i,j+1}) & \text{if } \beta_{i,j+\frac{1}{2}} = \beta_{i,j} = \beta_{i,j+1} = 0, \\ \tilde{h}_{i,j+\frac{1}{2}} + (\mathbf{M}_{\mathbf{L}}^{\mathbf{y},\text{FV}} \cdot \tilde{\eta}_{i,j} + \mathbf{M}_{\mathbf{R}}^{\mathbf{y},\text{FV}} \cdot \tilde{\eta}_{i,j+1}) & \text{if } \beta_{i,j+\frac{1}{2}} = \beta_{i,j} = \beta_{i,j+1} = 1, \\ \tilde{h}_{i,j+\frac{1}{2}} + (\mathbf{M}_{\mathbf{L}}^{\mathbf{y},\text{FV}} \cdot \mathcal{P}^{xy} \cdot \tilde{\eta}_{i,j} + \mathbf{M}_{\mathbf{R}}^{\mathbf{y},\text{FV}} \cdot \tilde{\eta}_{i,j+1}) & \text{if } \beta_{i,j+\frac{1}{2}} = \beta_{i,j+1} = 1, \beta_{i,j} = 0, \\ \tilde{h}_{i,j+\frac{1}{2}} + (\mathbf{M}_{\mathbf{L}}^{\mathbf{y},\text{FV}} \cdot \tilde{\eta}_{i,j} + \mathbf{M}_{\mathbf{R}}^{\mathbf{y},\text{FV}} \cdot \mathcal{P}^{xy} \cdot \tilde{\eta}_{i,j+1}) & \text{if } \beta_{i,j+\frac{1}{2}} = \beta_{i,j} = 1, \beta_{i,j+1} = 0. \end{cases} \quad (4.114)$$

The projection of the momentum  $U$  and  $V$  in  $x$  and  $y$  direction onto the main grid reads

$$\tilde{U}_{i,j} = \begin{cases} \mathbf{M}^{-1} \cdot (\mathbf{M}_{\mathbf{L}}^{\mathbf{x},\text{DG}} \cdot \tilde{U}_{i-\frac{1}{2},j} + \mathbf{M}_{\mathbf{R}}^{\mathbf{x},\text{DG}} \cdot \tilde{U}_{i+\frac{1}{2},j}) & \text{if } \beta_{i,j} = \beta_{i-\frac{1}{2},j} = \beta_{i+\frac{1}{2},j} = 0, \\ \mathbf{M}_{\mathbf{L}}^{\mathbf{x},\text{FV}} \cdot \tilde{U}_{i-\frac{1}{2},j} + \mathbf{M}_{\mathbf{R}}^{\mathbf{x},\text{FV}} \cdot \tilde{U}_{i+\frac{1}{2},j} & \text{if } \beta_{i,j} = \beta_{i-\frac{1}{2},j} = \beta_{i+\frac{1}{2},j} = 1, \\ \mathbf{M}^{-1} \cdot (\mathbf{M}_{\mathbf{L}}^{\mathbf{x},\text{DG}} \cdot \tilde{U}_{i-\frac{1}{2},j} + \mathbf{M}_{\mathbf{R}}^{\mathbf{x},\text{DG}} \cdot \mathcal{W}_{\mathcal{L}}^x \cdot \tilde{U}_{i+\frac{1}{2},j}) & \text{if } \beta_{i,j} = \beta_{i-\frac{1}{2},j} = 0, \beta_{i+\frac{1}{2},j} = 1, \\ \mathbf{M}^{-1} \cdot (\mathbf{M}_{\mathbf{L}}^{\mathbf{x},\text{DG}} \cdot \mathcal{W}_{\mathcal{R}}^x \cdot \tilde{U}_{i-\frac{1}{2},j} + \mathbf{M}_{\mathbf{R}}^{\mathbf{x},\text{DG}} \cdot \tilde{U}_{i+\frac{1}{2},j}) & \text{if } \beta_{i,j} = \beta_{i+\frac{1}{2},j} = 0, \beta_{i-\frac{1}{2},j} = 1, \end{cases} \quad (4.115)$$



$$\tilde{V}_{i,j} = \begin{cases} M^{-1} \cdot (M_L^{y,DG} \cdot \tilde{V}_{i,j-\frac{1}{2}} + M_R^{y,DG} \cdot \tilde{V}_{i,j+\frac{1}{2}}) \\ \text{if } \beta_{i,j} = \beta_{i,j-\frac{1}{2}} = \beta_{i,j+\frac{1}{2}} = 0, \\ M_L^{y,FV} \cdot \tilde{V}_{i,j-\frac{1}{2}} + M_R^{y,FV} \cdot \tilde{V}_{i,j+\frac{1}{2}} \\ \text{if } \beta_{i,j} = \beta_{i,j-\frac{1}{2}} = \beta_{i,j+\frac{1}{2}} = 1, \\ M^{-1} \cdot (M_L^{y,DG} \cdot \tilde{V}_{i,j-\frac{1}{2}} + M_R^{y,DG} \cdot \mathcal{W}_{\mathcal{L}}^y \cdot \tilde{V}_{i,j+\frac{1}{2}}) \\ \text{if } \beta_{i,j} = \beta_{i,j-\frac{1}{2}} = 0, \beta_{i,j+\frac{1}{2}} = 1, \\ M^{-1} \cdot (M_L^{y,DG} \cdot \mathcal{W}_{\mathcal{R}}^y \cdot \tilde{V}_{i,j-\frac{1}{2}} + M_R^{y,DG} \cdot \tilde{V}_{i,j+\frac{1}{2}}) \\ \text{if } \beta_{i,j} = \beta_{i,j+\frac{1}{2}} = 0, \beta_{i,j-\frac{1}{2}} = 1. \end{cases} \quad (4.116)$$

With the momentum averaged on the main grid it is possible to compute the nonlinear convective terms of the shallow water system. In the first MOOD step they are evaluated with the RKDG method shown in sub-section 4.2.2 while during the second MOOD step the convective terms are recomputed with the explicit finite volume method, see the end of section 4.2.3, but only in the control volumes flagged as troubled cells. Finally, is it easy to see that the stability conditions in eq. (4.79) and in eq. (4.90) are equivalent so the time step  $\Delta t$  for the limited case does not change with respect the one of the unlimited semi-implicit DG scheme. This completes the description of the two-dimensional algorithm.

### 4.3 Numerical tests for the one-dimensional model

In the following, the new staggered semi-implicit DG schemes with *a posteriori* subcell finite volume limiting are validated by comparing the numerical results against different reference solutions available in literature. In order to show the benefits of the new limiter, we present numerical results obtained with and without using the limiter. In all cases presented here, the initial condition has not been smoothed out, unlike in [54]. For the sake of clarity, we recall that when we explicitly write the value of the implicitness parameter  $\theta_{DG}$  this is referred to the zones where the limiter is not activated, while in the limited cells we always impose  $\theta = 1$  for the subcell finite volume limiter. Note also that during the first MOOD step  $\theta = \theta_{DG}$  is constant in the entire domain.

#### 4.3.1 Well-balanced property for the 1D case

We simulate the numerical test proposed in [106] in order to check numerically the well-balanced property (or C-property) of the limited DG scheme. We recall that a well-balanced method preserves exactly steady state solutions of the type  $\eta = \text{const}$  and  $u = 0$  even for a not-flat bottom topography, see [10, 80, 81, 83, 106]. The semi-implicit DG method is well balanced by construction, see [54] for details. Then, for the semi-implicit sub-cell finite volume scheme it is easy to see from eq. (4.24) that if the free surface is constant and the discharge was initially zero then it remains zero. This property is valid also for the combined DG-FV scheme. In order to verify the well balancing of the scheme also numerically, we introduce a computational domain  $\Omega_x = [-2, 2]$  and the bottom topography is a smooth profile given as follows

$$h(x) = \begin{cases} -0.25(\cos(10\pi(x - 1.5)) + 1) & \text{if } 1.4 \leq x \leq 1.6 \\ 0 & \text{otherwise.} \end{cases} \quad (4.117)$$

First, we impose a flat free surface  $\eta = 1$  and zero initial velocity and we run a simulation up to the final time  $t = 0.1$  using  $N_x = 100$ ,  $P = 5$ ,  $\theta_{DG} = 0.55$  and  $\Delta t_0 = 10^{-3}$ . Moreover, since for a smooth problem the MOOD approach would not detect any troubled cells and since we want to verify the well-balanced property of the scheme with active subcell FV limiter, the limiter is *randomly* activated in each time step in some of the control volumes. At the end of the simulation, we observe that the free surface is still flat and in Table 4.1 we expose the results for the  $L_\infty$  and  $L_2$  norms referred to the quantity  $\eta$  computed using different machine precisions. In all the cases the errors have the same magnitudes of the machine tolerance, confirming the well-balanced property of the numerical method. Similar results are also obtained for the fluid velocity  $u$ .

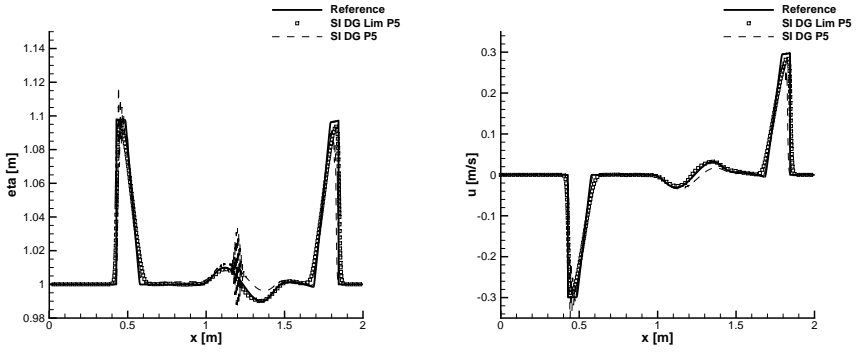
Case	$L_\infty$	$L_2$
Single precision	$2.2245 \cdot 10^{-6}$	$8.7232 \cdot 10^{-6}$
Double precision	$1.6088 \cdot 10^{-13}$	$6.2094 \cdot 10^{-13}$
Quadruple precision	$4.4248 \cdot 10^{-31}$	$1.7787 \cdot 10^{-30}$

**Table 4.1:** Numerical validation of the well-balanced property for the semi-implicit staggered DG scheme in 1D with sub-cell FV limiter.  $L_\infty$  and  $L_2$  norms for  $\eta$  with  $P = 5$  at time  $t = 0.1$  for different machine precisions.

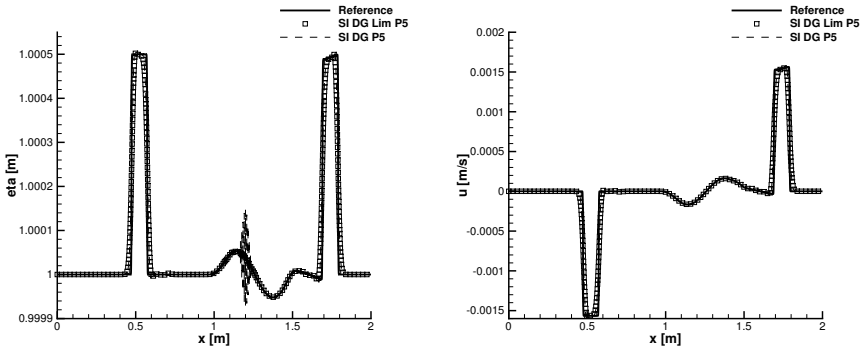
Now we impose a discontinuous perturbation in the free surface elevation of the following form

$$\eta(x, 0) = \begin{cases} 1 + \epsilon & \text{if } 1.1 \leq x \leq 1.2 \\ 1 & \text{otherwise,} \end{cases} \quad (4.118)$$

and we consider other two test cases: for the first one we set  $\epsilon = 0.2$  while for the second one  $\epsilon = 10^{-3}$ . For both simulations we use  $N_x = 250$  and a polynomial degree of the basis functions of  $P = 5$ . The final time of the simulations is  $t = 0.2$ , with a time step size of  $\Delta t_0 = 4 \cdot 10^4$  and  $\theta_{DG} = 0.55$ . The results are shown in Figs. 4.19 and 4.20. For this test a reference solution is produced using a robust explicit second order TVD finite volume scheme with a well balanced Osher Riemann solver [63]. It is possible to see a good agreement between the reference solution and the numerical one given by the limited DG scheme. The *a posteriori* limiter is active only in the first time steps of the simulation, while there are no troubled cells detected at the final time. However, the limiting in the first time steps is necessary, since the unlimited DG scheme produces visible spurious oscillations in the regions close to the discontinuities in the free surface profile, but also in the central region where the reference solution is smooth. We attribute this behaviour to the fact that the unlimited DG scheme is not viscous enough for cleaning the Gibbs phenomenon that is caused by the discontinuities in the initial condition. We emphasize again that in the limited DG scheme, the limiter acts only in the first time steps of the simulations, but successfully avoids the generation and propagation of unphysical oscillations.



**Figure 4.19:** Reference solutions and numerical solutions for the LeVeque test with  $\epsilon = 0.2$  at  $t = 0.2$ .



**Figure 4.20:** Reference solutions and numerical solutions for the LeVeque test with  $\epsilon = 10^{-3}$  at  $t = 0.2$ .

### 4.3.2 One-dimensional Riemann problems

Now we expose the numerical numerical results for a set of *Riemann problems*. The Riemann problem is a typical benchmark in the context of hyperbolic PDE and it consists in a discontinuous initial condition of the state vector  $\mathbf{Q} = [\eta, u, h]$  of the type,

$$\mathbf{Q}(x, 0) = \begin{cases} \mathbf{Q}_L & \text{if } x \leq 0 \\ \mathbf{Q}_R & \text{if } x > 0, \end{cases} \quad (4.119)$$

i.e. two piecewise constant states, separated by a discontinuity. The number of cells  $N_x$  used to discretize the computational domain  $\Omega_x = [x_L, x_R]$  is chosen so that  $\Delta x = 0.1333$  and the polynomial degree is chosen as  $P = 5$ . For all the simulations the parameters are  $\Delta t_0 = 5 \cdot 10^{-4}$  and  $\theta_{DG} = 0.55$ . We consider eight different initial conditions (see Table 4.2) and the exact solution is given by the Riemann solver of Toro [147] in the case of flat bottom and the one of Bernetti et al. [11] when the bottom is discontinuous.

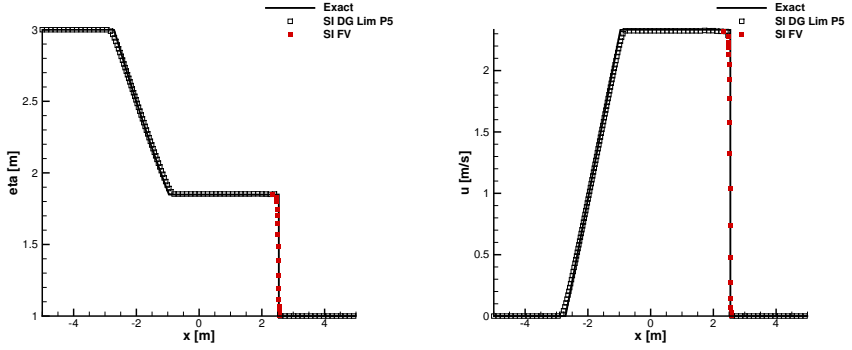
The results of RP1 are exposed in Fig. 4.21 and the test consists in a shock wave that travels to the right and a rarefaction wave that moves to the left. The numerical solution fits the exact one very well and the troubled cells are located around the shock wave, as expected, while the plateau and the rarefaction wave are properly resolved by the unlimited DG scheme. It is also interesting to note that it was impossible to run this test case without any limiting of the semi-implicit staggered DG scheme, because at a certain point the spurious oscillations produce negative water depth and then the simulation crashes.

In the test RP2 two flows collide and they generate two shocks that travel in opposite direction. From Fig. 4.22 we observe that the troubled cells are detected in a perfectly symmetric way and the staggered semi-implicit DG scheme with a *a posteriori* subcell FV limiter reproduces properly the exact solution, activating the semi-implicit finite volume solver only in the troubled zones around the shock waves. On the contrary, the pure unlimited DG scheme produces evident Gibbs phenomena that are not acceptable.

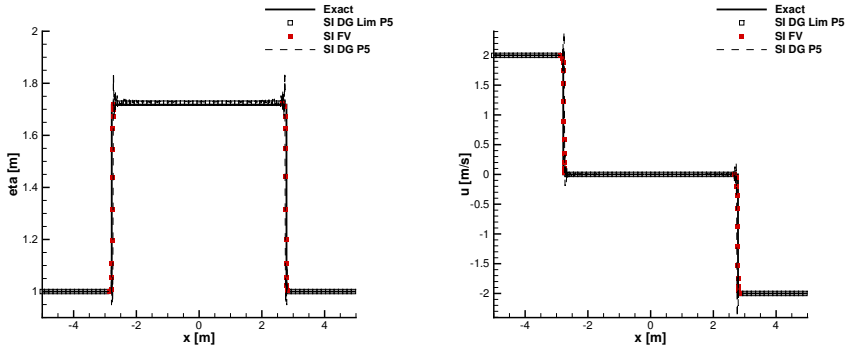
The test cases RP3, RP4 and RP5 are taken from [54]. At the final time of the simulation the semi-implicit DG scheme with a *a posteriori* subcell FV limiter gives a very good match with the exact solution and troubled cells are not detected. However, it is possible that advancing in time some troubled cells are detected and thus evolved by the more robust finite volume method for a few time steps, especially in the regions characterized by the presence of shock waves. In general, this behaviour happens using a rather high polynomial degree and a parameter  $\theta$  that tends to 0.5. Moreover, the unlimited semi-implicit DG scheme produces solutions characterized by unphysical oscillations close to the discontinuities and even in the flat zones of the Riemann problem.

The Riemann problems RP6 (from [54]) and RP7 (from [11]) are characterized by a discontinuity in the bottom topography. Also in these cases the agreement with the reference solution is good for the DG scheme with subcell limiter, while the unlimited DG scheme crashes in the last two test problems.

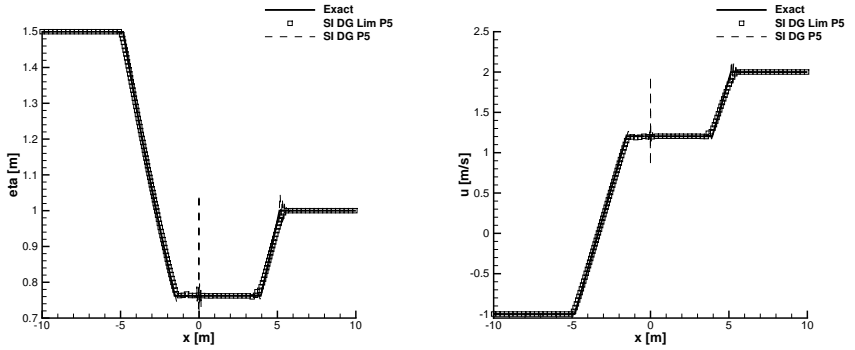
In the last two configurations we consider Riemann problems which have been solved by using the new algorithm for wetting and drying presented in subsection 4.1.6. RP8 is a classical dam break problem over dry bed while in RP9 the initial state is a discontinuity in the velocity profile that produces two rarefactions and a dry region in the middle of the channel. In both cases the agreement with the reference solutions is very good and these results are shown in figures 4.28 and 4.29. In conclusion, the semi-implicit staggered DG scheme with sub-cell FV limiter is in good agreement with the reference solution even in the presence of discontinuities. Furthermore, it significantly improves the quality of the numerical results with respect to the unlimited case.



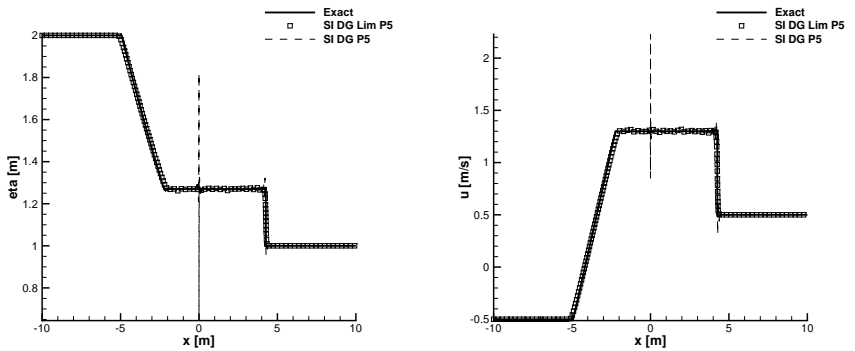
**Figure 4.21:** Reference solutions and numerical solutions for the Riemann problem RP1 at  $t = 1.0$ .



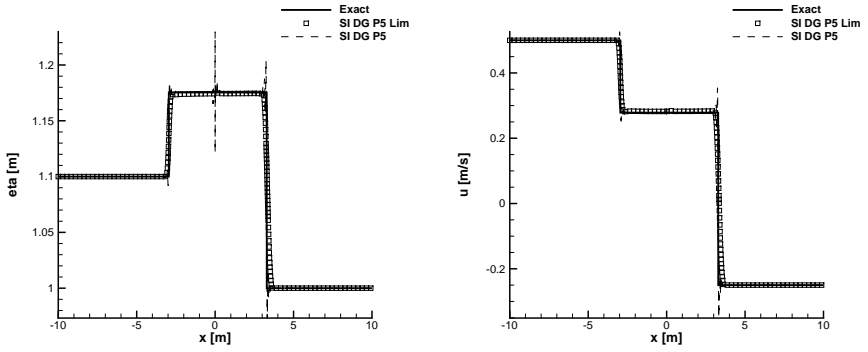
**Figure 4.22:** Reference solutions and numerical solutions for the Riemann problem RP2 at  $t = 1.0$ .



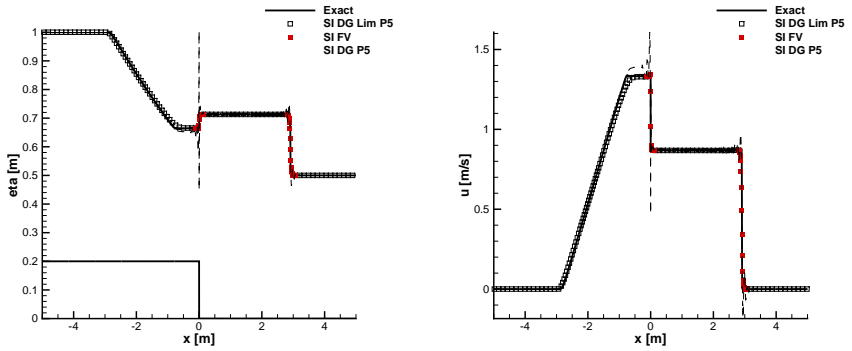
**Figure 4.23:** Reference solutions and numerical solutions for the Riemann problem RP3 at  $t = 1.0$ .



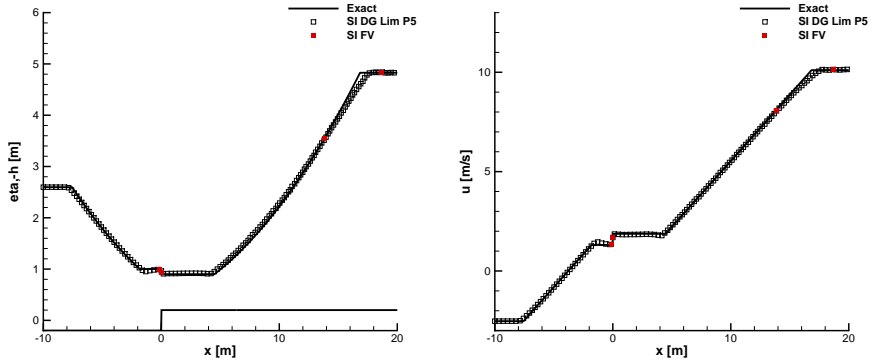
**Figure 4.24:** Reference solutions and numerical solutions for the Riemann problem RP4 at  $t = 1.0$ .



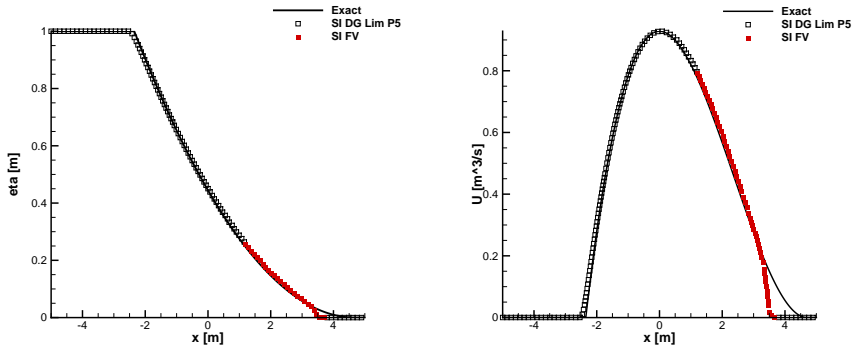
**Figure 4.25:** Reference solutions and numerical solutions for the Riemann problem RP5 at  $t = 1.0$ .



**Figure 4.26:** Reference solutions and numerical solutions for the Riemann problem RP6 at  $t = 1.0$ .



**Figure 4.27:** Reference solutions and numerical solutions for the Riemann problem RP7 at  $t = 1.0$ .

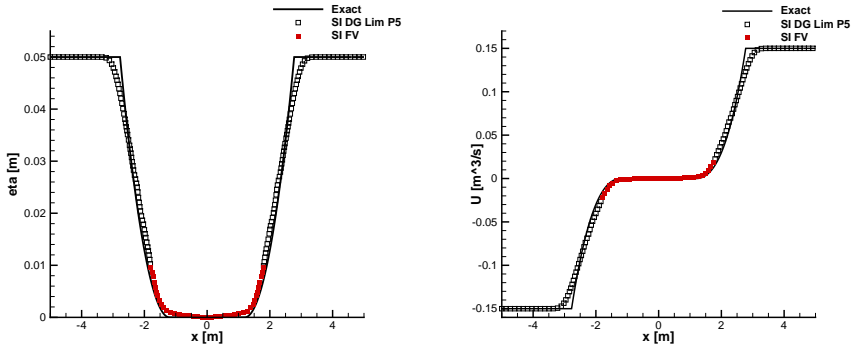


**Figure 4.28:** Reference solutions and numerical solutions for the Riemann problem RP8 at  $t = 0.75$ .



Case	$\eta_L$	$u_L$	$h_L$	$\eta_R$	$u_R$	$h_R$	$x_L$	$x_R$	$N_x$	$t_{end}$
RP1	3.0	0.0	0.0	1.0	0.0	0.0	-5.0	+5.0	75	1.0
RP2	1.0	2.0	0.0	1.0	-2.0	0.0	-5.0	+5.0	75	1.0
RP3	1.5	-1.0	0.0	1.0	+2.0	0.0	-10.0	+10.0	150	1.0
RP4	2.0	-0.5	0.0	1.0	+0.5	0.0	-10.0	+10.0	150	1.0
RP5	1.1	0.5	0.0	1.0	-0.25	0.0	-10.0	+10.0	150	1.0
RP6	1.0	0.0	-0.2	0.5	0.0	0.0	-10.0	+10.0	150	1.0
RP7	2.6	-2.5	0.0	4.8	10.1	-0.2	-10.0	+20.0	225	1.0
RP8	1.0	0.0	0.0	0.0	0.0	0.0	-5.0	+5.0	50	0.75
RP9	0.05	-3.0	0.0	0.05	+3.0	0.0	-5.0	+5.0	50	0.7

**Table 4.2:** Riemann problems for semi-implicit DG on staggered grid with sub-cell limiter - Left and right states for  $\eta, u$  and  $h$ , left and right boundaries  $x_L$  and  $x_R$ , number of control volume  $N_x$ ,  $\Delta_x$  and final time of the simulations  $t_{end}$ .



**Figure 4.29:** Reference solutions and numerical solutions for the Riemann problem RP9 at  $t = 0.7$ .

### 4.3.3 Trans-critical flow over a bump

For this 1D test case, the bottom geometry is given by the following function, which describes a smooth obstacle centred in  $x = 0$

$$h(x) = \begin{cases} -0.25(\cos(10\pi(x-1)) + 1) & \text{if } -0.1 \leq x \leq 0.1 \\ 0 & \text{otherwise} \end{cases} \quad (4.120)$$

and a constant initial state is imposed for the free surface and for the velocity:  $\eta = 1$  and  $u = 0.3\sqrt{g}$ . The computational domain is  $\Omega_x = [x_L, x_R] = [-2.5, 2.5]$  and is covered by 300 cells setting  $P = 7$ . The implicitness parameter is set to  $\theta_{DG} = 1$  because at the final time of the simulation  $t_{end} = 0.5$  the free surface profile is in a steady regime in the region over the bump. The maximum time step is  $\Delta t_0 = 5 \cdot 10^{-4}$ . We expose the steady state profile in the domain  $[-1, +1]$  and we report the numerical solution for  $\eta$ ,  $u$ ,  $H$  and the Froude number  $Fr = u/\sqrt{gH}$ . The results in Fig. 4.30 show an excellent agreement between the numerical results obtained with the staggered semi-implicit DG scheme with subcell FV limiter and the reference solution, which has been again obtained with a second order TVD finite volume scheme using the path-conservative Osher-type Riemann solver presented in [63]. Over the bump we observe a supercritical flow characterized by a very high Froude number. Moreover, we underline that for this set of numerical parameters, the discontinuity and the hydraulic jump are contained in the same troubled cell. Also for this case we observed that without limiter the simulation does not reach the final time.

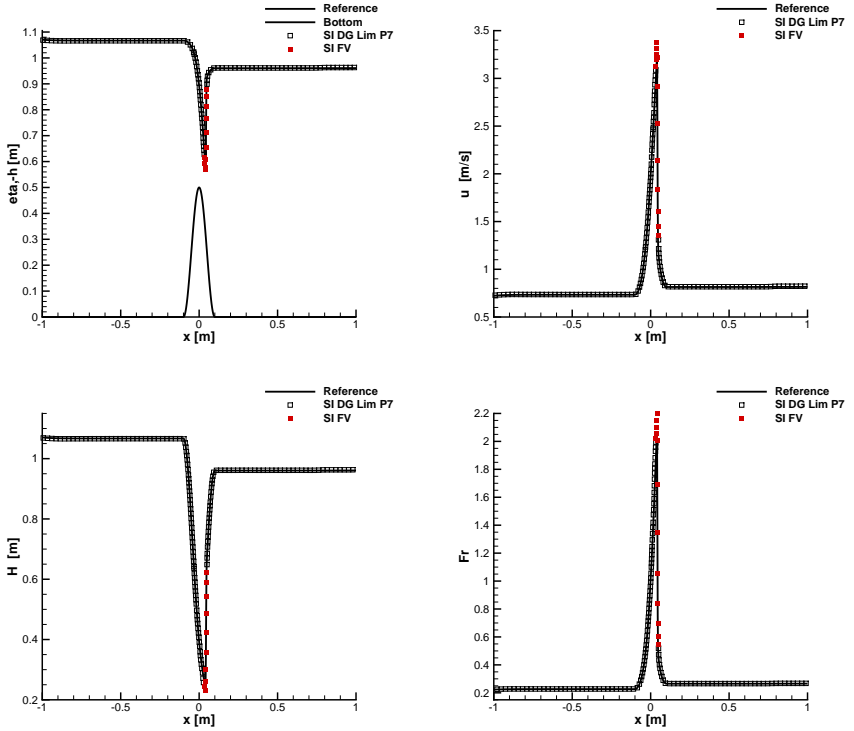
### 4.3.4 Oscillating lake in a parabolic bowl

In this last 1D benchmark problem we consider a periodic flow with wetting and drying front in a channel with parabolic bottom profile. The computational domain is  $\Omega_x = [x_L, x_R] = [-2L, 2L]$ , where  $L = 2500$  is the characteristic length of the problem. The bottom is prescribed as  $h(x) = -D_0 L^{-1} x^2$  and the analytical solutions for the free-surface and for the velocity are ([113, 144])

$$\eta(x, t) = \begin{cases} D_0 + \frac{2AD_0}{L^2}(x - \frac{A}{2}) & \text{if } A \cos(\omega t) - L \leq x \leq A \cos(\omega t) + L, \\ \frac{D_0}{L} x^2 & \text{otherwise,} \end{cases} \quad (4.121)$$

$$u(x, t) = \begin{cases} -A\omega \cos(\omega t) & \text{if } A \cos(\omega t) - L \leq x \leq A \cos(\omega t) + L, \\ 0 & \text{otherwise,} \end{cases} \quad (4.122)$$

where  $D_0 = 10$  is the maximum equilibrium depth,  $A = 0.5L$  denotes the amplitude of the oscillation and  $\omega = \sqrt{2gD_0}L^{-2}$  is the angular frequency. For this simulation we use  $N_x = 100$  with a polynomial degree  $P = 5$  and  $\theta_{DG} = 0.6$ . The numerical data for the free surface  $\eta$  and for the discharge  $Q = U = uH$  are compared against the exact solution in figures 4.31, 4.32, 4.33 and 4.34 at time  $\tau/4$ ,  $\tau/2$ ,  $3\tau/4$  and  $\tau$  respectively, with  $\tau = 2\pi/\omega$ . One can observe that in these figures the agreement between numerical and analytical solution is excellent.



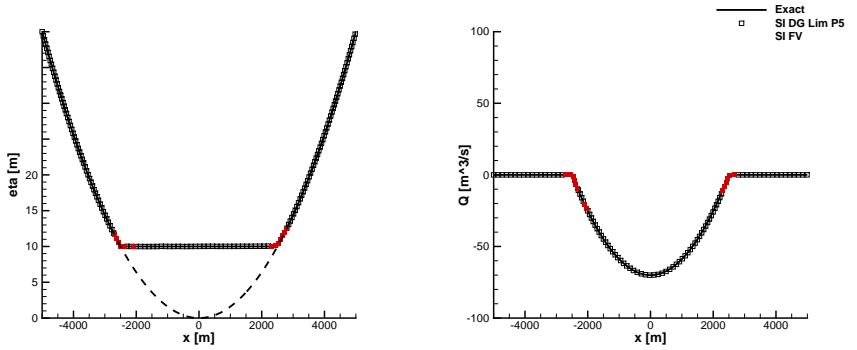
**Figure 4.30:** Reference solution and numerical solution for the trans-critical flow over a bump at  $t = 1.0$ .

## 4.4 Numerical tests for the two-dimensional model

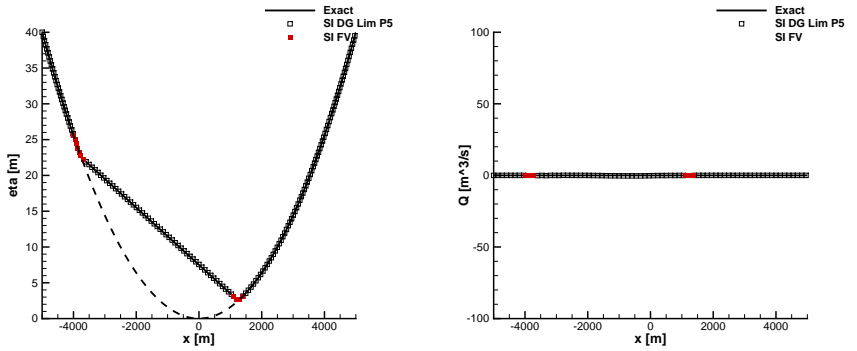
### 4.4.1 Well-balanced property for the 2D case

Here we verify the C-property of the semi-implicit staggered DG scheme with subcell FV limiter also for the 2D case. For this purpose, a smooth bottom profile is defined as

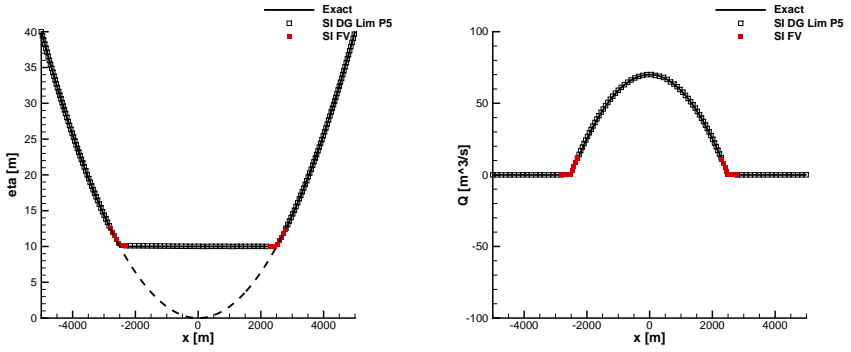
$$h(x, y) = -0.8e^{-5(x-0.9)^2-50(y-0.5)^2}. \quad (4.123)$$



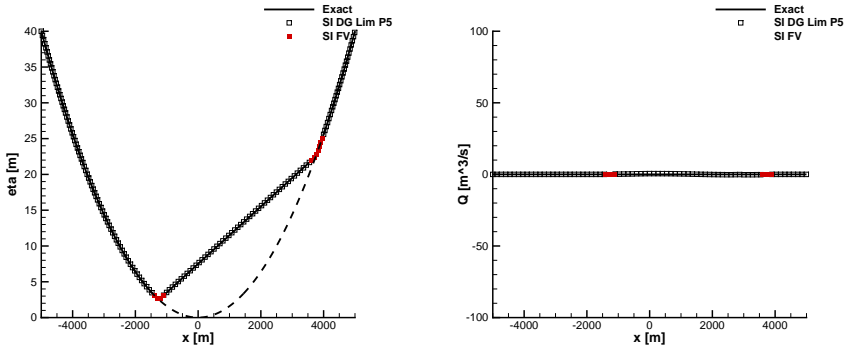
**Figure 4.31:** Oscillating lake - Reference solutions and numerical solutions for  $\eta$  and  $Q$  at  $t = \tau/4$ .



**Figure 4.32:** Oscillating lake - Reference solutions and numerical solutions for  $\eta$  and  $Q$  at  $t = \tau/2$ .



**Figure 4.33:** Oscillating lake - Reference solutions and numerical solutions for  $\eta$  and  $Q$  at  $t = 3\tau/4$ .



**Figure 4.34:** Oscillating lake - Reference solutions and numerical solutions for  $\eta$  and  $Q$  at  $t = \tau$ .

First, a flat free surface  $\eta(x, y, 0) = 1$  is imposed at time  $t = 0$  and the velocity is initially set to zero. The computational domain  $\Omega_{xy} = [0, 2] \times [0, 1]$  is paved with  $80 \times 40$  cells in  $x$  and  $y$  direction, respectively. The polynomial degree is  $P = 5$ ,  $\Delta t_0 = 10^{-3}$ ,  $\theta_{DG} = 0.55$  and the final time is  $t = 0.1$ . Similarly to the 1D case, the limiter status flag  $\beta$  is randomly assigned a value of either 0 or 1. As in the 1D case, the error measured in different norms is very close to the chosen machine precision, see Table 4.4.1.

Case	$L_\infty$	$L_2$
Single precision	$6.7949 \cdot 10^{-6}$	$7.4372 \cdot 10^{-7}$
Double precision	$3.4972 \cdot 10^{-15}$	$3.7973 \cdot 10^{-16}$
Quadruple precision	$4.4970 \cdot 10^{-31}$	$4.9041 \cdot 10^{-31}$

**Table 4.3:** Numerical validation of the well-balanced property for the semi-implicit staggered DG scheme with subcell FV limiter in 2D.  $L_\infty$  and  $L_2$  norms for  $\eta$  with  $P = 5$  at time  $t = 0.1$  for different machine precisions.

For the last test, the following discontinuous free-surface profile is assigned as initial condition

$$\eta(x, y, 0) = \begin{cases} 1 + \epsilon & \text{if } 0.05 \leq x \leq 0.15 \\ 1 & \text{otherwise.} \end{cases} \quad (4.124)$$

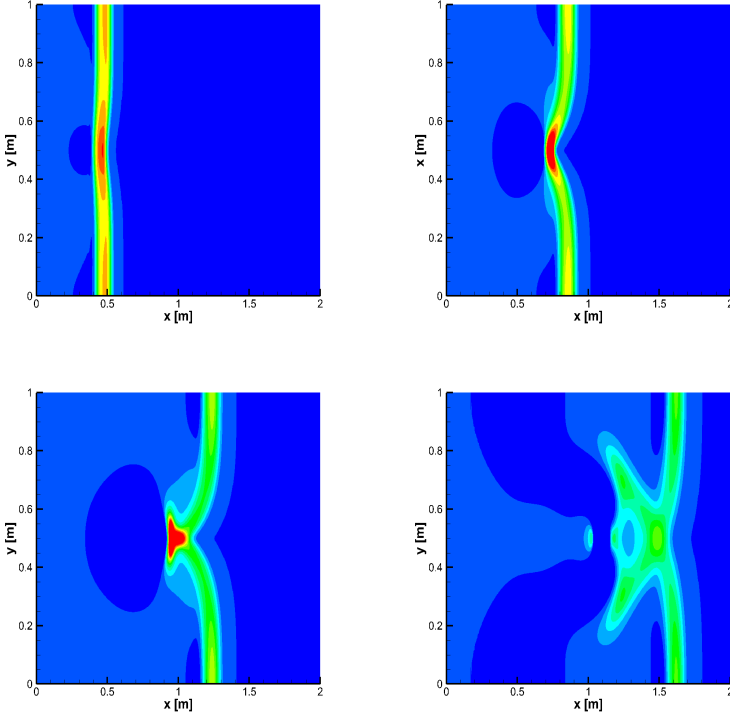
We use  $N_x = 120$  and  $N_y = 40$  with  $P = 5$  on the domain  $\Omega_{xy} = [-1, 2] \times [0, 1]$ . The implicitness parameter is set to  $\theta_{DG} = 0.55$  and  $\Delta t_0 = 10^{-3}$ . In Fig. 4.35 we show the color contours of the free surface elevation  $\eta$  at different output times. We observe that the limiter is activated only during the first time steps of the simulation and then the simulation produces smooth outputs, similar to those presented in [106, 136].

#### 4.4.2 Circular dambreak

Finally, we consider two circular dambreak problems of the following form

$$\mathbf{Q}(x, y, 0) = \begin{cases} \mathbf{Q}_{\text{in}} & \text{if } r \leq r_c \\ \mathbf{Q}_{\text{out}} & \text{if } r > r_c. \end{cases} \quad (4.125)$$

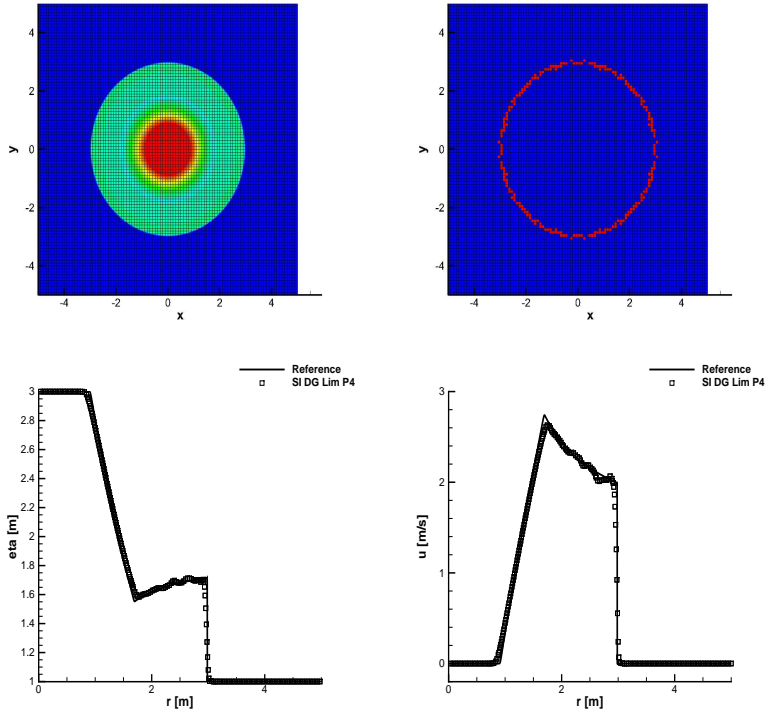
where  $r = \sqrt{x^2 + y^2}$  and  $\mathbf{Q}$  is the vector of state defined as  $\mathbf{Q} = [h, u, v, h]$ . In the first case the bottom geometry is a flat surface and the inner and outer states are given by  $\mathbf{Q}_{\text{in}} = [3, 0, 0, 0]$  and  $\mathbf{Q}_{\text{out}} = [1, 0, 0, 0]$  with  $r_c = 2$ . The computational domain is  $\Omega_{xy} = [-5, +5] \times [-5, +5]$  and we use  $N_x = N_y = 100$  with  $\Delta t_0 = 10^{-3}$ ,  $\theta_{DG} = 0.55$  and  $t_{\text{end}} = 0.2$ . For the second test an additional discontinuity is introduced in order to create a jump in the bottom, setting  $\mathbf{Q}_{\text{in}} = [2, 0, 0, -0.2]$  and  $\mathbf{Q}_{\text{out}} = [1, 0, 0, 0]$  with  $r_c = 1$ . The computational domain is  $\Omega_{xy} = [-2, +2] \times [-2, +2]$  and also in this case  $N_x = N_y = 100$  with  $\Delta t_0 = 10^{-3}$ ,



**Figure 4.35:** Numerical validation of the well-balanced property for the 2D case. Numerical solution of the free surface elevation  $\eta$  at times  $t = 0.12$ ,  $t = 0.24$ ,  $t = 0.36$  and  $t = 0.48$  from top left to bottom right.

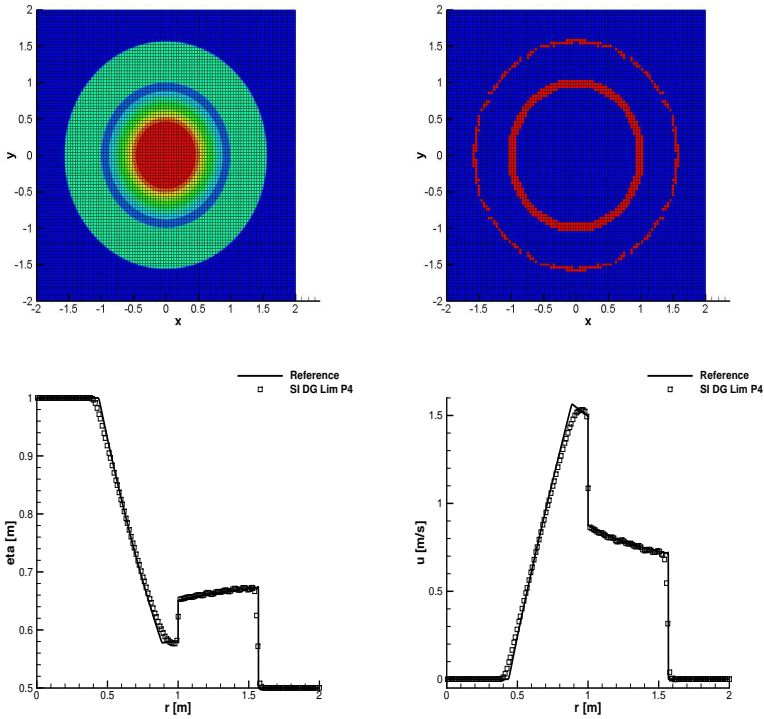
$\theta_{DG} = 0.55$  and  $t_{end} = 0.2$ . In both test cases we set the polynomial degree to  $P = 5$ . In the top rows of Figs. 4.36 and 4.37 we present the numerical results for the free surface elevation as well as the colour maps of the troubled zones highlighted by the red cells, while unlimited cells are plotted in blue. The results show that the FV limiter is properly activated at shocks and discontinuities only. In the bottom rows of Figs. 4.36 and 4.37 we compare the numerical results for  $\eta$  and  $u$  against the reference solution that was computed by solving an equivalent 1D problem in radial direction with

geometric source terms using a second order TVD finite volume scheme on a very fine grid, see [147] for details. In both cases the results are satisfactory and the semi-implicit DG scheme with *a posteriori* subcell FV limiter produces only some small dispersive oscillations.



**Figure 4.36:** Circular dam break problem with flat bottom. Contour plot of the free surface (Top left) and map of the troubled cells coloured in red (Top right); 1D cut along the  $x$  axis comparing the reference solution with the numerical solution for the free surface  $\eta$  (Bottom left) and for the velocity  $u$  (Bottom right).





**Figure 4.37:** Circular dam break problem with discontinuous bottom. Contour plot of the free surface (Top left) and map of the troubled cells coloured in red (Top right); 1D cut along the  $x$  axis comparing the reference solution with the numerical solution for the free surface  $\eta$  (Bottom left) and for the velocity  $u$  (Bottom right).

## 4.5 Conclusions about the new subcell finite volume limiter for staggered semi-implicit DG schemes applied to the shallow water equations

In this PhD thesis we have extended the *a posteriori* sub-cell finite volume limiter method introduced in [14, 59, 64] for explicit DG schemes on collocated grids to semi-implicit discontinuous Galerkin methods on staggered meshes. The new limiter has been applied to the solution of the shallow water equations in one and two space dimensions. To our knowledge, this is the first time that the *a posteriori* subcell FV limiting strategy has been applied to semi-implicit DG schemes on staggered grids. This novel family of methods is based on the high order MOOD paradigm introduced in [39, 47, 48, 107] in the finite volume context. First, the unlimited semi-implicit staggered DG scheme presented in [54] is applied in order to get the *candidate solution* at time  $t^{n+1}$ , which is obtained by solving a linear (or mildly nonlinear) system of equations for the free surface elevation. In the next step, *troubled cells* which are characterized by a non-admissible solution are found thanks to physical and numerical admissibility criteria. In control volumes marked as troubled cells, then a more robust semi-implicit staggered finite volume scheme is applied and, consequently, the linear (or mildly nonlinear) system for the free surface elevation has to be solved again, this time involving the unlimited DG cells and the finite volume sub-cells.

Several test problems have been solved and they confirm that the new method works well and is able to suppress spurious oscillations near flow discontinuities. In particular, a sequence of Riemann problems has been solved, demonstrating a significant improvement in the quality of the final results compared to the unlimited DG scheme. In fact, only very few unphysical oscillations are produced because the limiter is able to stabilize the DG scheme in a nonlinear (data-dependent) manner, activating the subcell finite volume method only in those zones where the detector finds troubled cells. Via numerical experiments we have also shown that the new scheme is able to deal with wet and dry fronts properly, thanks to the nonlinear wetting and drying algorithm of Casulli [28] used within the *a posteriori* subcell finite volume limiter.

## 5 A posteriori sub-cell finite volume limiting of staggered SIDG schemes for the compressible Euler equations

In this chapter we propose a novel semi-implicit Discontinuous Galerkin (DG) scheme on staggered meshes with *a posteriori* sub-cell finite volume limiting for the one and two dimensional Euler equations of compressible gas dynamics, extending the strategy that was outlined in the previous chapter for the shallow water equations. We follow the strategy adopted by Dumbser and Casulli in 2016 where the Euler equations have been solved by using a semi-implicit finite volume method based on the flux-vector splitting proposed by Toro and Vázquez-Cendón [152]. In particular, the non-linear convective terms are discretized explicitly and then the pressure is discretized implicitly and it is obtained by solving a linear system. As consequence, the time step is given by a mild CFL condition based only on the fluid velocity and not based on the sound speed, which makes this new method also suitable for simulations in the low Mach number regime.

In addition, in order to deal with shock waves or strong discontinuities, the scheme includes the *a posteriori* sub-cell finite volume limiting technique. This strategy was proposed by Dumbser et al. in 2014 for explicit DG schemes and it based on the MOOD algorithm of Clain, Loubère and Diot. This approach was extended to semi-implicit DG scheme on staggered meshes for the shallow water equations in Chapter 4. In particular, first, at time  $t^n$  the unlimited DG scheme produces a so-called candidate solution for the time level  $t^{n+1}$ . Later on, the control volumes with a non-admissible candidate solution are identified by using physical and numerical detection criteria in order to check the positivity of the solution, the absence of floating point errors and the respect of a relaxed discrete maximum principle (DMP). Then a robust, stable in the sense of Godunov, first order semi-implicit finite volume (FV) method is applied on a sub-grid composed of  $2P + 1$  cells, where  $P$  denotes the polynomial approximation degree used in the DG scheme. Successively, the nonlinear convective terms are updated by the explicit subcell finite volume scheme, while the linear system for the new pressure is reassembled and solved again, using the DG method in unlimited cells and the subcell FV method in troubled cells. Finally, in all troubled cells the higher order DG polynomials are reconstructed again from the piecewise constant subcell finite volume averages and the scheme proceeds with the next time step.

We validate this novel family of methods and we carry out some classical numerical benchmarks of gas dynamics. Great attention is dedicated to 1D and 2D Riemann problems and we show that for these test cases the scheme works well in the presence of shock waves and it does not produce non-physical oscillations.

## 5.1 Staggered semi-implicit DG schemes with a posteriori sub-cell finite volume limiting for the Euler equations in 1D

In this section we introduce the novel staggered semi-implicit DG schemes with a posteriori sub-cell finite volume limiting for the 1D Euler equations. For the sake of simplicity, we first derive the unlimited semi-implicit 1D DG scheme in 5.1.2; then in 5.1.3 we discuss a semi-implicit FV method with a proper subcell formulation. Later on, we recall the MOOD algorithm and we conclude in 5.1.5 with the general scheme that combines the two methods previously mentioned. In the following, several universal tensors will be introduced and for clarity we report them apart in Appendix A.3.

### 5.1.1 Governing equations of the 1D model

The Euler equations of gas dynamics in one space dimension are a well-known system of hyperbolic PDEs and read as follows:

$$\begin{aligned} \frac{\partial \rho}{\partial t} + \frac{\partial(\rho u)}{\partial x} &= 0, \\ \frac{\partial(\rho u)}{\partial t} + \frac{\partial}{\partial x}(\rho u^2 + p) &= 0, \\ \frac{\partial(\rho E)}{\partial t} + \frac{\partial}{\partial x}(u(\rho E + p)) &= 0. \end{aligned} \tag{5.1}$$

The equations (5.1) state the principles of mass, momentum and total energy conservation. The spatial and the temporal coordinates are  $x$  and  $t$  and the computational domain is denoted by  $\Omega_x = [x_L, x_R]$  in the following. Furthermore,  $\rho(x, t)$  is the fluid density,  $u(x, t)$  is the velocity,  $p(x, t)$  is the pressure and  $E(x, t)$  is the specific total energy. Additionally, we have  $E = e + k$ , where  $e$  is the specific internal energy and  $k = \frac{1}{2}u^2$  is the specific kinetic energy. We introduce another quantity called specific enthalpy  $h = e + p/\rho$  and consequently we rewrite the flux of the energy equation as the sum of two different contributions  $u(\rho E + p) = u\rho k + h\rho u$ , see [55, 152]. In order to close the system we use the equation of state (EOS) for an ideal gas, which reads

$$e = e(p, \rho) = \frac{p}{(\gamma - 1)\rho}. \tag{5.2}$$

Here,  $\gamma = c_p/c_v$  is the ratio of specific heats, which typically lies in the range  $1 < \gamma < 3$ . Here we will consider this quantity as a constant and equal to  $\gamma = 1.4$ , which is the adiabatic index of a diatomic gas and thus a reasonable value for air at moderate pressures and temperatures.

The system in eq. (5.1) can be written in more compact matrix-vector notation as

$$\frac{\partial \mathbf{Q}}{\partial t} + \frac{\partial \mathbf{F}(\mathbf{Q})}{\partial x} = 0, \tag{5.3}$$

where  $\mathbf{Q} = [\rho, \rho u, \rho E]^T$  is the vector of conserved variables and  $\mathbf{F} = [\rho u, \rho u^2 + p, u(\rho E + p)]^T$  is the physical flux vector. In quasi-linear form the above system reads

$$\frac{\partial \mathbf{Q}}{\partial t} + \mathbf{A}(\mathbf{Q}) \frac{\partial \mathbf{Q}}{\partial x} = 0, \tag{5.4}$$

where  $\mathbf{A}(\mathbf{Q}) = \partial \mathbf{F}(\mathbf{Q}) / \partial \mathbf{Q}$  is the Jacobian matrix of the system. For the Euler equations,  $\mathbf{A}(\mathbf{Q})$  has three real eigenvalues:  $\lambda_1 = u - a$ ,  $\lambda_2 = u$  and  $\lambda_3 = u + a$ , where  $a$  is the *sound speed*. For an ideal gas the sound speed is equal to  $a = \sqrt{\gamma p / \rho}$ . Moreover, the matrix  $\mathbf{A}$  has a complete set of linearly independent eigenvectors, hence the system is hyperbolic. For a thorough discussion of the Euler equations of gas dynamics and numerical methods for their discretization, the reader is referred to the textbook of Toro [148]. Following the flux vector splitting proposed by Toro and Vázquez-Cendón in [152] and the ideas outlined in [55, 66], the PDE system (5.3) can now be rewritten as

$$\frac{\partial \mathbf{Q}}{\partial t} + \frac{\partial \mathbf{F}^c(\mathbf{Q})}{\partial x} + \frac{\partial \mathbf{F}^p(\mathbf{Q})}{\partial x} = 0, \quad (5.5)$$

with the nonlinear convective flux  $\mathbf{F}^c(\mathbf{Q}) = [\rho u, \rho u^2, u \rho k]^T$  that does not contain any pressure terms, and the pressure flux  $\mathbf{F}^p(\mathbf{Q}) = [0, p, h \rho u]^T$ . Following [55] and [66], we will discretize  $\mathbf{F}^c$  explicitly and  $\mathbf{F}^p$  implicitly. In [152] it was shown that the Jacobian of the convective flux  $\mathbf{A}^c(\mathbf{Q}) = \partial \mathbf{F}^c(\mathbf{Q}) / \partial \mathbf{Q}$  has the eigenvalues  $\lambda_1^c = 0$  and  $\lambda_{2,3}^c = u$ , i.e. the CFL condition on the explicit part of the scheme will depend only on the bulk flow velocity  $u$  and not on the sound speed  $a$ . For this reason, the method presented here can also be used for the simulation of low Mach number flows.

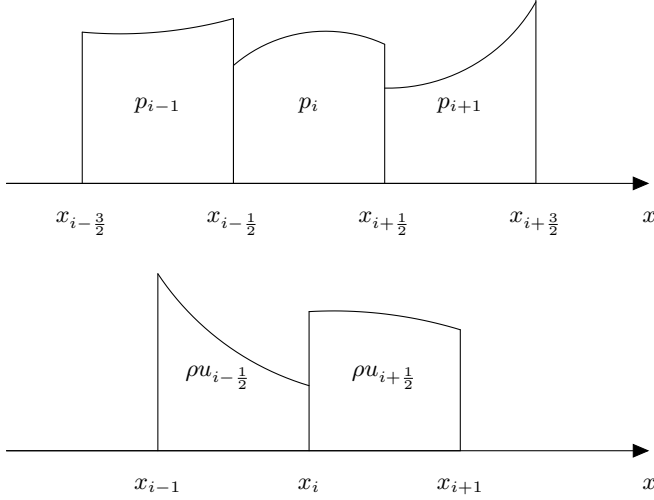
### 5.1.2 Unlimited staggered semi-implicit DG scheme for the 1D Euler equations

We consider a computational domain  $\Omega_x = [x_L, x_R]$  composed of two overlapping grids (see Fig. 5.1). The first one is called main grid, which contains  $N_x$  cells characterized by a constant length  $\Delta x = \frac{L}{N_x} = \frac{x_R - x_L}{N_x}$ , while the second grid is called dual mesh and it has  $N_x + 1$  elements. On this last grid there are  $N_x - 1$  equally spaced elements of size  $\Delta x$ , while the length of the cells on the left and on the right boundary is only  $\Delta x/2$ . A generic element of the main grid is denoted by  $T_i = [x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}]$ , whereas a control volume on the staggered grid is indicated as  $T_{i+\frac{1}{2}} = [x_i, x_{i+1}]$  and its edges are the barycenters of two consecutive cells of the main grid. In order to represent our numerical solution via piecewise polynomials of degree  $P$ , we use a set of basis functions  $\varphi(\xi)$  defined on the reference element  $[0, 1]$ . In particular, we choose a *nodal basis* given by the Lagrange interpolation polynomials passing through the Gauss-Legendre quadrature points of the reference element  $[0, 1]$ . This basis is by construction orthogonal. For any physical control volume we generate the basis functions  $\phi_l(x)$  on the main grid and the basis functions  $\psi_l(x)$  on the dual grid from  $\varphi_l(\xi)$  as

$$\begin{aligned} \phi_l(x) &= \varphi_l(\xi) \quad \text{with} \quad x = x_i + \xi \Delta x, \quad \text{and} \\ \psi_l(x) &= \varphi_l(\xi) \quad \text{with} \quad x = x_{i+\frac{1}{2}} + \xi \Delta x, \quad 0 \leq \xi \leq 1. \end{aligned} \quad (5.6)$$

Then, a quantity located on the main grid, for example the pressure, is approximated as follows

$$p_i(x, t) = \sum_{i=1}^{P+1} \phi_l(x) \hat{p}_{i,l}(t) := \boldsymbol{\phi}(\mathbf{x}) \cdot \hat{\mathbf{p}}_i(t), \quad (5.7)$$



**Figure 5.1:** Staggered grids for the one-dimensional DG scheme. Main grid used for the free surface (top) and dual mesh for the velocity (bottom).

while for variables that are defined on the staggered dual mesh, such as the fluid velocity, we have

$$u_{i+\frac{1}{2}}(x, t) = \sum_{l=1}^{P+1} \psi_l(x) \hat{u}_{i+\frac{1}{2}, l}(t) := \boldsymbol{\psi}(x) \cdot \hat{\mathbf{u}}_{i+\frac{1}{2}}(t) \quad (5.8)$$

where  $\hat{\mathbf{p}}$  and  $\hat{\mathbf{u}}$  are the degrees of freedom for the pressure and for the velocity, respectively. Following the ideas of semi-implicit methods, see e.g. [55], the Euler equations can be split into a convective sub-system, which contains the non linear terms and into a linear pressure subsystem.

### 5.1.2.1 Explicit discretization of the nonlinear convective terms

In the numerical scheme proposed in this chapter, the convective subsystem

$$\frac{\partial \mathbf{Q}}{\partial t} + \frac{\partial \mathbf{F}^c(\mathbf{Q})}{\partial x} = 0 \quad (5.9)$$

is integrated using an *explicit* Runge-Kutta discontinuous Galerkin (RKDG) scheme on the main grid, based on a third order TVD Runge-Kutta method [82]. For this purpose, we define the discrete

solution for the vector  $\mathbf{Q}$  within an element  $T_i$  as

$$\mathbf{Q}_i(t) = \phi_l(x) \hat{\mathbf{Q}}_{i,l}(t), \quad (5.10)$$

where we assume the Einstein summation convention over repeated indices. Multiplication of eq. (5.9) with a test function  $\phi_k$ , integration over the primary control volume  $T_i$  with subsequent integration by parts and introduction of a numerical flux lead to the following semi-discrete method

$$\int_{T_i} \phi_k(x) \frac{\partial \mathbf{Q}_i}{\partial t} dx + \phi_k(x_{i+\frac{1}{2}}) \mathbf{F}_{i+\frac{1}{2}}^c - \phi_k(x_{i-\frac{1}{2}}) \mathbf{F}_{i-\frac{1}{2}}^c - \int_{T_i} \frac{\partial \phi_k}{\partial x} \mathbf{F}^c(\mathbf{Q}_i) dx = 0. \quad (5.11)$$

The numerical fluxes  $\mathbf{F}_{i\pm\frac{1}{2}}^c$  at the cell interfaces  $x_{i\pm\frac{1}{2}}$  are given by a local Lax-Friedrichs (Rusanov) Riemann solver

$$\mathbf{F}_{i+\frac{1}{2}}^c = \frac{1}{2} \left( \mathbf{F}^c \left( \mathbf{Q}_{i+\frac{1}{2}}^+ \right) + \mathbf{F}^c \left( \mathbf{Q}_{i+\frac{1}{2}}^- \right) \right) - \frac{1}{2} s_{\max} \left( \mathbf{Q}_{i+\frac{1}{2}}^+ - \mathbf{Q}_{i+\frac{1}{2}}^- \right), \quad (5.12)$$

where  $\mathbf{Q}_{i+\frac{1}{2}}^\pm$  are the boundary extrapolated values at the element interfaces from the right and left, respectively. Since in this explicit part of the scheme we only consider the convective subsystem, the maximal signal speed  $s_{\max}$  is given by the eigenvalues  $\lambda^c$  of the convective flux  $\mathbf{F}^c$ , i.e.

$$s_{\max} = \max \left( \max \left| \lambda^c \left( \mathbf{Q}_{i+\frac{1}{2}}^+ \right) \right|, \max \left| \lambda^c \left( \mathbf{Q}_{i+\frac{1}{2}}^- \right) \right| \right) = \max \left( |u_{i+\frac{1}{2}}^+|, |u_{i+\frac{1}{2}}^-| \right). \quad (5.13)$$

After integration with a third order TVD Runge-Kutta scheme [82] the degrees of freedom of the conservative variables of the convective subsystem at the new time are denoted by  $\hat{\mathbf{Q}}^*$  and read as follows<sup>1</sup>:

$$\begin{aligned} \hat{\mathbf{k}}_1 &= \hat{\mathbf{Q}}^n + \Delta t \mathbf{L}_h \left( \hat{\mathbf{Q}}^n \right), \\ \hat{\mathbf{k}}_2 &= \frac{3}{4} \hat{\mathbf{Q}}^n + \frac{1}{4} \hat{\mathbf{k}}_1 + \frac{1}{4} \Delta t \mathbf{L}_h \left( \hat{\mathbf{k}}_1 \right), \\ \hat{\mathbf{Q}}^* &= \frac{1}{3} \hat{\mathbf{Q}}^n + \frac{2}{3} \hat{\mathbf{k}}_2 + \frac{2}{3} \Delta t \mathbf{L}_h \left( \hat{\mathbf{k}}_2 \right). \end{aligned} \quad (5.14)$$

In eq. (5.14) the spatial discretization operator  $\mathbf{L}_h \left( \hat{\mathbf{Q}}^n \right)$  is defined for each element  $T_i$  as

$$\mathbf{L}_h \left( \hat{\mathbf{Q}} \right) \Big|_{T_i} = -\frac{1}{\Delta x} \mathbf{M}^{-1} \left( \varphi(1) \mathbf{F}_{i+\frac{1}{2}}^c - \varphi(0) \mathbf{F}_{i-\frac{1}{2}}^c - \int_0^1 \varphi'(\xi) \varphi(\xi) d\xi \cdot \hat{\mathbf{F}}_i^c \right), \quad (5.15)$$

with the element mass matrix on the reference element given by

$$\mathbf{M} = \int_0^1 \varphi \varphi d\xi = \int_0^1 \varphi_k \varphi_l d\xi, \quad (5.16)$$

and the vector of degrees of freedom of the convective flux simply defined as  $\hat{\mathbf{F}}_i^c = \mathbf{F}^c \left( \hat{\mathbf{Q}}_i \right)$ , since we use an orthogonal nodal basis.

<sup>1</sup> when the index  $i$  is omitted in the vector of degrees of freedom we intend the entire set of all degrees of freedom of all elements

### 5.1.2.2 Derivation of the semi-implicit staggered DG scheme in 1D

We first discretize the total energy equation obtained after the flux splitting procedure (5.5). We multiply the total energy equation by the vector of test functions  $\phi$  and we integrate over a control volume of the main grid  $T_i$ :

$$\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \phi \left( \frac{\partial}{\partial t} (\rho e + \rho k) + \frac{\partial}{\partial x} (\rho k u + h \rho u) \right) dx = 0. \quad (5.17)$$

In the time derivative, the term of the total energy term at  $t^{n+1}$  is divided considering the contribution of the specific internal energy and the one of the kinetic energy; in addition we make use of the term  $\widehat{\rho E}_i^*$  that has been obtained after the explicit discretization of the nonlinear convective terms discussed before. Then, we integrate the spatial derivative by parts in space and, in order to achieve second order of accuracy in time, we use the so-called  $\theta$ -method. This leads to

$$\begin{aligned} & \left( \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \phi \phi dx \right) \cdot \left( \frac{\widehat{\rho e}_i^{n+1} + \widehat{\rho k}_i^{n+1} - \widehat{\rho E}_i^*}{\Delta t} \right) \\ & + \phi(x_{i+\frac{1}{2}}^-) \psi(x_{i+\frac{1}{2}}) \psi(x_{i+\frac{1}{2}}) \cdot \hat{h}_{i+\frac{1}{2}}^{n+\theta_{i+\frac{1}{2}}} \widehat{\rho u}_{i+\frac{1}{2}}^{n+\theta_{i+\frac{1}{2}}} \\ & - \phi(x_{i-\frac{1}{2}}^+) \psi(x_{i-\frac{1}{2}}) \psi(x_{i-\frac{1}{2}}) \cdot \hat{h}_{i-\frac{1}{2}}^{n+\theta_{i-\frac{1}{2}}} \widehat{\rho u}_{i-\frac{1}{2}}^{n+\theta_{i-\frac{1}{2}}} \\ & - \left( \int_{x_{i-\frac{1}{2}}}^{x_i} \frac{\partial \phi}{\partial x} \psi \psi dx \right) \cdot \hat{h}_{i-\frac{1}{2}}^{n+\theta_{i-\frac{1}{2}}} \widehat{\rho u}_{i-\frac{1}{2}}^{n+\theta_{i-\frac{1}{2}}} - \left( \int_{x_i}^{x_{i+\frac{1}{2}}} \frac{\partial \phi}{\partial x} \psi \psi dx \right) \cdot \hat{h}_{i+\frac{1}{2}}^{n+\theta_{i+\frac{1}{2}}} \widehat{\rho u}_{i+\frac{1}{2}}^{n+\theta_{i+\frac{1}{2}}} \\ & = 0, \end{aligned} \quad (5.18)$$

where  $\widehat{\rho u}_{i+\frac{1}{2}}^{n+\theta_{i+\frac{1}{2}}} = (1 - \theta_{i+\frac{1}{2}}) \widehat{\rho u}_{i+\frac{1}{2}}^n + \theta_{i+\frac{1}{2}} \widehat{\rho u}_{i+\frac{1}{2}}^{n+1}$  and  $0.5 \leq \theta \leq 1$  is an implicitness parameter. When  $\theta = 1$  the scheme is the implicit Euler scheme which is first order accurate in time; when  $\theta = 0.5$  we have the second-order Crank-Nicolson time discretization [32]. According to the Godunov theorem [78], in this last case the method is not monotone and can generate unphysical oscillations in the vicinity of strong discontinuities. As in the previous Chapter 4, in the framework of sub-cell FV limiters for semi-implicit DG schemes we therefore want to vary the parameter  $\theta$  in the domain and this is done assigning different values of  $\theta$  on the dual mesh. Eq (5.18) can be rewritten in the compact matrix-vector form

$$\mathbf{M} \cdot (\widehat{\rho e}_i^{n+1} + \widehat{\rho k}_i^{n+1} - \widehat{\rho E}_i^*) + \frac{\Delta t}{\Delta x} (\mathbf{R}_e^{\text{DG}} \cdot \hat{h}_{i+\frac{1}{2}}^{n+\theta_{i+\frac{1}{2}}} \widehat{\rho u}_{i+\frac{1}{2}}^{n+\theta_{i+\frac{1}{2}}} - \mathbf{L}_e^{\text{DG}} \cdot \hat{h}_{i-\frac{1}{2}}^{n+\theta_{i-\frac{1}{2}}} \widehat{\rho u}_{i-\frac{1}{2}}^{n+\theta_{i-\frac{1}{2}}}) = 0. \quad (5.19)$$



where  $\mathbf{R}_e^{\text{DG}} \cdot \hat{\mathbf{h}}_{i+\frac{1}{2}}^{n+\theta} \widehat{\rho \mathbf{u}}_{i+\frac{1}{2}}^{n+\theta}$  is the product of a rank 3 tensor  $\mathbf{R}_e^{\text{DG}}$  with the degrees of freedom  $\hat{\mathbf{h}}_{i+\frac{1}{2}}^{n+\theta} \widehat{\rho \mathbf{u}}_{i+\frac{1}{2}}^{n+\theta}$ . Moving the quantities of eq. (5.19) at time  $t^n$  to the right hand side yields

$$\begin{aligned} & \mathbf{M} \cdot \frac{\widehat{\mathbf{p}}_i^{n+1}}{\gamma-1} + \mathbf{M} \cdot \widehat{\rho \mathbf{k}}_i^{n+1} + \frac{\Delta t}{\Delta x} (\theta_{i+\frac{1}{2}} \mathbf{R}_e^{\text{DG}} \cdot \hat{\mathbf{h}}_{i+\frac{1}{2}}^{n+1} \widehat{\rho \mathbf{u}}_{i+\frac{1}{2}}^{n+1} - \theta_{i-\frac{1}{2}} \mathbf{L}_e^{\text{DG}} \cdot \hat{\mathbf{h}}_{i-\frac{1}{2}}^{n+1} \widehat{\rho \mathbf{u}}_{i-\frac{1}{2}}^{n+1}) = \\ & \mathbf{M} \cdot \widehat{\rho \mathbf{E}}_i^* - \frac{\Delta t}{\Delta x} ((1-\theta_{i+\frac{1}{2}}) \mathbf{R}_e^{\text{DG}} \cdot \hat{\mathbf{h}}_{i+\frac{1}{2}}^n \widehat{\rho \mathbf{u}}_{i+\frac{1}{2}}^n - (1-\theta_{i-\frac{1}{2}}) \mathbf{L}_e^{\text{DG}} \cdot \hat{\mathbf{h}}_{i-\frac{1}{2}}^n \widehat{\rho \mathbf{u}}_{i-\frac{1}{2}}^n) \end{aligned} \quad (5.20)$$

where we expressed the internal energy density  $\rho e$  as function of  $p$  using the ideal gas equation of state. In addition, the meaning of the symbol under-tilde will be clarified later. The momentum equation is multiplied by the test function  $\psi$  and it is integrated over a control volume  $T_{i+\frac{1}{2}}$  that belongs to the dual grid

$$\int_{x_i}^{x_{i+1}} \psi \left( \frac{\partial \rho u}{\partial t} + \frac{\partial \rho u^2}{\partial x} \right) dx = - \int_{x_i}^{x_{i+1}} \psi \frac{\partial p}{\partial x} dx. \quad (5.21)$$

We carry out the integration of (5.21) following the same procedure used in [54, 90, 92, 140]. The nonlinear convective terms have already been discretized explicitly, leading to  $\widehat{\rho \mathbf{u}}_{i+\frac{1}{2}}^*$ , while the pressure gradient is approximated by splitting it into a smooth contribution and a jump term across the interfaces of the primary grid, which are located in the interior of the dual mesh. This yields

$$\begin{aligned} & \left( \int_{x_i}^{x_{i+1}} \psi \psi dx \right) \cdot \left( \frac{\widehat{\rho \mathbf{u}}_{i+\frac{1}{2}}^{n+1} - \widehat{\rho \mathbf{u}}_{i+\frac{1}{2}}^*}{\Delta t} \right) = \\ & - \psi(x_{i+\frac{1}{2}}) \left( \phi(x_{i+\frac{1}{2}}^+) \cdot \hat{\mathbf{p}}_{i+1}^{n+\theta} - \phi(x_{i+\frac{1}{2}}^-) \cdot \hat{\mathbf{p}}_i^{n+\theta} \right) \\ & - \left( \int_{x_i}^{x_{i+\frac{1}{2}}^-} \psi \frac{\partial \phi}{\partial x} dx \right) \cdot \hat{\mathbf{p}}_i^{n+\theta} - \left( \int_{x_{i+\frac{1}{2}}^+}^{x_{i+1}} \psi \frac{\partial \phi}{\partial x} dx \right) \cdot \hat{\mathbf{p}}_{i+1}^{n+\theta} \end{aligned} \quad (5.22)$$

and the more compact matrix-tensor formulation of the same equation reads as follows:

$$\mathbf{M} \cdot (\widehat{\rho \mathbf{u}}_{i+\frac{1}{2}}^{n+1} - \widehat{\rho \mathbf{u}}_{i+\frac{1}{2}}^*) + \frac{\Delta t}{\Delta x} \left( \mathbf{R}_p^{\text{DG}} \cdot \hat{\mathbf{p}}_{i+1}^{n+\theta} - \mathbf{L}_p^{\text{DG}} \cdot \hat{\mathbf{p}}_i^{n+\theta} \right) = 0. \quad (5.23)$$

Later on, we introduce the term  $\widehat{\mathbf{G}}_{i+\frac{1}{2}}^n$  that collects the convective terms of the momentum equation and the discrete pressure gradient at the time level  $t^n$

$$\widehat{\mathbf{G}}_{i+\frac{1}{2}}^n = \widehat{\rho \mathbf{u}}_{i+\frac{1}{2}}^* - \frac{\Delta t}{\Delta x} (1-\theta_{i+\frac{1}{2}}) \mathbf{M}^{-1} \cdot (\mathbf{R}_p^{\text{DG}} \cdot \hat{\mathbf{p}}_{i+1}^n - \mathbf{L}_p^{\text{DG}} \cdot \hat{\mathbf{p}}_i^n), \quad (5.24)$$

so that the final semi-implicit momentum equation reads

$$\widehat{\rho u}_{i+\frac{1}{2}}^{n+1} = \widehat{G}_{i+\frac{1}{2}}^n - \frac{\Delta t}{\Delta x} \theta_{i+\frac{1}{2}} \mathbf{M}^{-1} \cdot (\mathbf{R}_{\mathbf{p}}^{\text{DG}} \cdot \widehat{\mathbf{p}}_{i+\frac{1}{2}}^{n+1} - \mathbf{L}_{\mathbf{p}}^{\text{DG}} \cdot \widehat{\mathbf{p}}_i^{n+1}). \quad (5.25)$$

Later, we will insert the discrete momentum equation (5.25) into the discrete total energy equation (5.20). However, the enthalpy  $h$  is a function of the pressure, and in order to avoid a strongly non-linear system in  $p$  we adopt an iterative *Picard* technique. Hence, the symbol under-tilde means that a variable is evaluated at the previous Picard iteration. In the following we will use the superscript  $r$  as the iteration index of the Picard process. For the sake of clarity we rewrite eq. (5.20) as

$$\begin{aligned} \mathbf{M} \cdot \frac{\widehat{\mathbf{p}}_i^{n+1,r+1}}{\gamma - 1} + \frac{\Delta t}{\Delta x} (\theta_{i+\frac{1}{2}} \mathbf{R}_{\mathbf{e}}^{\text{DG}} \cdot \widehat{\mathbf{h}}_{i+\frac{1}{2}}^{n+1,r} \widehat{\rho u}_{i+\frac{1}{2}}^{n+1,r+1} - \theta_{i-\frac{1}{2}} \mathbf{L}_{\mathbf{e}}^{\text{DG}} \cdot \widehat{\mathbf{h}}_{i-\frac{1}{2}}^{n+1,r} \widehat{\rho u}_{i-\frac{1}{2}}^{n+1,r+1}) = \\ \mathbf{M} \cdot (\widehat{\rho E}_i^* - \widehat{\rho k}_i^{n+1,r}) \\ - \frac{\Delta t}{\Delta x} ((1 - \theta_{i+\frac{1}{2}}) \mathbf{R}_{\mathbf{e}}^{\text{DG}} \cdot \widehat{\mathbf{h}}_{i+\frac{1}{2}}^n \widehat{\rho u}_{i+\frac{1}{2}}^n - (1 - \theta_{i-\frac{1}{2}}) \mathbf{L}_{\mathbf{e}}^{\text{DG}} \cdot \widehat{\mathbf{h}}_{i-\frac{1}{2}}^n \widehat{\rho u}_{i-\frac{1}{2}}^n). \end{aligned} \quad (5.26)$$

The discrete momentum equation (5.25) becomes

$$\widehat{\rho u}_{i+\frac{1}{2}}^{n+1,r+1} = \widehat{G}_{i+\frac{1}{2}}^n - \frac{\Delta t}{\Delta x} \theta_{i+\frac{1}{2}} \mathbf{M}^{-1} \cdot (\mathbf{R}_{\mathbf{p}}^{\text{DG}} \cdot \widehat{\mathbf{p}}_{i+\frac{1}{2}}^{n+1,r+1} - \mathbf{L}_{\mathbf{p}}^{\text{DG}} \cdot \widehat{\mathbf{p}}_i^{n+1,r+1}). \quad (5.27)$$

Therefore, after inserting (5.27) into (5.26) one gets a discrete wave equation for the unknown pressure  $p^{n+1,r+1}$  at the next Picard iteration, which can be solved by using a Thomas algorithm for linear block three-diagonal systems, and which reads

$$\mathcal{L}_i^{\text{DG}} \cdot \widehat{\mathbf{p}}_{i-1}^{n+1,r+1} + \mathcal{C}_i^{\text{DG}} \cdot \widehat{\mathbf{p}}_i^{n+1,r+1} + \mathcal{R}_i^{\text{DG}} \cdot \widehat{\mathbf{p}}_{i+1}^{n+1,r+1} = \widehat{\mathbf{b}}_i^r. \quad (5.28)$$

Here,  $\widehat{\mathbf{b}}_i^r$  is the known right hand side term

$$\begin{aligned} \widehat{\mathbf{b}}_i^r = \mathbf{M} \cdot (\widehat{\rho E}_i^* - \widehat{\rho k}_i^{n+1,r}) \\ - \frac{\Delta t}{\Delta x} (\theta_{i+\frac{1}{2}} \mathbf{R}_{\mathbf{e}}^{\text{DG}} \cdot \widehat{\mathbf{h}}_{i+\frac{1}{2}}^{n+1,r} \widehat{\rho u}_{i+\frac{1}{2}}^* - \theta_{i-\frac{1}{2}} \mathbf{L}_{\mathbf{e}}^{\text{DG}} \cdot \widehat{\mathbf{h}}_{i-\frac{1}{2}}^{n+1,r} \widehat{\rho u}_{i-\frac{1}{2}}^*) \\ - \frac{\Delta t}{\Delta x} ((1 - \theta_{i+\frac{1}{2}}) \mathbf{R}_{\mathbf{e}}^{\text{DG}} \cdot \widehat{\mathbf{h}}_{i+\frac{1}{2}}^n \widehat{\rho u}_{i+\frac{1}{2}}^n - (1 - \theta_{i-\frac{1}{2}}) \mathbf{L}_{\mathbf{e}}^{\text{DG}} \cdot \widehat{\mathbf{h}}_{i-\frac{1}{2}}^n \widehat{\rho u}_{i-\frac{1}{2}}^n). \end{aligned} \quad (5.29)$$

The block matrices  $\mathcal{L}_i^{\text{DG}}$ ,  $\mathcal{C}_i^{\text{DG}}$  and  $\mathcal{R}_i^{\text{DG}}$  read

$$\mathcal{L}_i^{\text{DG}} = - \frac{\Delta t^2}{\Delta x^2} \theta_{i-\frac{1}{2}}^2 \mathbf{L}_{\mathbf{e}}^{\text{DG}} \widehat{\mathbf{h}}_{i-\frac{1}{2}}^{n+1,r} \mathbf{M}^{-1} \mathbf{L}_{\mathbf{p}}^{\text{DG}}, \quad (5.30)$$

$$\mathcal{C}_i^{\text{DG}} = \frac{1}{\gamma - 1} \mathbf{M} + \frac{\Delta t^2}{\Delta x^2} \theta_{i+\frac{1}{2}}^2 \mathbf{R}_{\mathbf{e}}^{\text{DG}} \widehat{\mathbf{h}}_{i+\frac{1}{2}}^{n+1,r} \mathbf{M}^{-1} \mathbf{L}_{\mathbf{p}}^{\text{DG}} + \frac{\Delta t^2}{\Delta x^2} \theta_{i-\frac{1}{2}}^2 \mathbf{L}_{\mathbf{e}}^{\text{DG}} \widehat{\mathbf{h}}_{i-\frac{1}{2}}^{n+1,r} \mathbf{M}^{-1} \mathbf{R}_{\mathbf{p}}^{\text{DG}}, \quad (5.31)$$

$$\mathcal{R}_i^{\text{DG}} = - \frac{\Delta t^2}{\Delta x^2} \theta_{i+\frac{1}{2}}^2 \mathbf{R}_{\mathbf{e}}^{\text{DG}} \widehat{\mathbf{h}}_{i+\frac{1}{2}}^{n+1,r} \mathbf{M}^{-1} \mathbf{R}_{\mathbf{p}}^{\text{DG}}. \quad (5.32)$$

For every Picard iteration, once the pressure is known, the momentum update is done using eq. (5.27). Subsequently, also the enthalpy  $\hat{h}_{i+\frac{1}{2}}^{n+1,r+1}$  has to be computed and, in order to get a quantity from the main grid to the dual one and back, we use the following averaging operator based on  $L_2$  projection:

$$\hat{p}_{i+\frac{1}{2}} = \mathbf{M}^{-1} \cdot (\mathbf{M}_L^{\text{DG}} \cdot \hat{p}_i + \mathbf{M}_R^{\text{DG}} \cdot \hat{p}_{i+1}). \quad (5.33)$$

In addition, since in eq. (5.29) the kinetic energy is needed on the main grid, we first compute the degrees of freedom of the velocity as  $\hat{u}_{i+\frac{1}{2}}^{n+1,r+1} = \widehat{\rho u}_{i+\frac{1}{2}}^{n+1,r+1} / \widehat{\rho}_{i+\frac{1}{2}}^{n+1}$  and then we apply the  $L_2$  averaging operator from the dual to the main grid

$$\hat{u}_i = \mathbf{M}^{-1} \cdot (\mathbf{M}_L^{\text{DG}} \cdot \hat{u}_{i-\frac{1}{2}} + \mathbf{M}_R^{\text{DG}} \cdot \hat{u}_{i+\frac{1}{2}}). \quad (5.34)$$

Numerical experiments indicate that only very few Picard iterations, typically 2 or 3, are enough to obtain a good solution. Finally, at the end of the Picard loop the update of the total energy is carried out and from eq. (5.19) we get

$$\widehat{\rho E}_i^{n+1} = \widehat{\rho E}_i^* - \frac{\Delta t}{\Delta x} \mathbf{M}^{-1} \cdot (\mathbf{R}_e^{\text{DG}} \cdot \hat{h}_{i+\frac{1}{2}}^{n+\theta} \widehat{\rho u}_{i+\frac{1}{2}}^{n+\theta} - \mathbf{L}_e^{\text{DG}} \cdot \hat{h}_{i-\frac{1}{2}}^{n+\theta} \widehat{\rho u}_{i-\frac{1}{2}}^{n+\theta}). \quad (5.35)$$

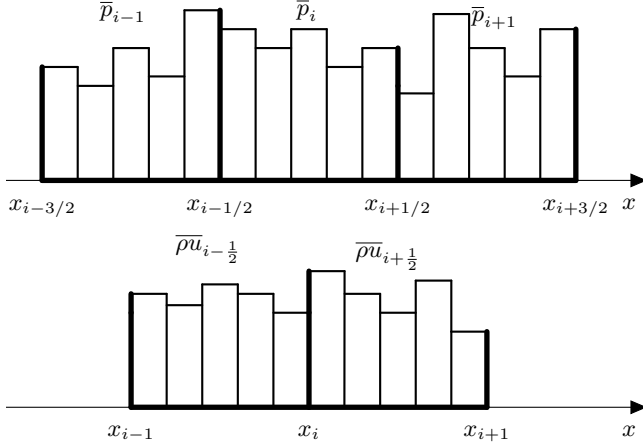
For this semi-implicit DG scheme, the maximum time step is given by the usual CFL condition of the RKDG scheme which reads

$$\Delta t < \text{CFL} \frac{\Delta x}{\max(|u|)} \quad \text{with} \quad \text{CFL} < (2P + 1)^{-1}. \quad (5.36)$$

Note that this is the stability condition due to the explicit discretization of the nonlinear convective terms and which is based only on the fluid velocity. Due to the implicit treatment of the pressure terms, the speed of sound does not influence the choice of  $\Delta t$ . Consequently, this method becomes suitable for the numerical simulation of low Mach number flows. However, due to its conservative formulation, the method is also able to deal with shock waves, as shown later. For other types of semi-implicit DG schemes on staggered meshes, the reader is referred to [54, 70, 71, 92, 137, 139, 140]. In addition, the scheme derived here represents the natural extension to arbitrary order of accuracy in space of the semi-implicit 1D finite volume method introduced in [55].

### 5.1.3 A sub-cell formulation for the semi-implicit finite volume method for the 1D Euler equations

In this subsection we introduce a sub-cell formulation of the semi-implicit finite volume method for the 1D Euler equations. In particular, this approach was adopted in Chapter 4 for the shallow water equations, but other similar methods have been proposed also in [27, 90, 99, 134]. The computational domain is  $\Omega_x = [x_L, x_R]$  and, similarly to the method in subsection 5.1.2, we consider two staggered meshes. For an element  $T_i$  of the main grid, given a positive integer  $P$ , there are  $2P + 1$  piecewise constant sub-cell averages that represent the FV data on  $2P + 1$  sub-cells  $T_{i,s}$  characterized by length  $\Delta x_s = \Delta x / (2P + 1)$ . Consequently we have that  $T_i = \cup_s T_{i,s}$  with  $s = 1, \dots, (2P + 1)$ . Hence, for  $P = 2$ ,  $\bar{p}_i = [\bar{p}_{i,1}, \bar{p}_{i,2}, \bar{p}_{i,3}, \bar{p}_{i,4}, \bar{p}_{i,5}]$  denotes the set of sub-cell data in the cell  $T_i$ . For a schematic representation of these grids see Fig. 5.2. The quantities assigned to main grid and to the dual grid are the same discussed in the previous section 5.1.2. For



**Figure 5.2:** Staggered grids for the sub-cell finite volume scheme in one dimension in the case  $P = 2$  ( $N_s = 5$ ). Main grid used for the free surface (top) and staggered dual mesh for the velocity (bottom).

the convective sub-system, the explicit finite volume update on a collocated grid at the subcell level is given by

$$\bar{Q}_{i,s}^* = \bar{Q}_{i,s}^n - \frac{\Delta t}{\Delta x_s} (\mathbf{F}_{i,s+\frac{1}{2}}^c - \mathbf{F}_{i,s-\frac{1}{2}}^c), \quad (5.37)$$

where we used again the Rusanov method in order to evaluate the numerical fluxes  $\mathbf{F}_{i,s\pm\frac{1}{2}}^c$ .

We start with the semi-implicit approximation of the momentum equation. Using the adaptive  $\theta$ -method, the discretization reads as follows

$$(\overline{\rho u}_{i+\frac{1}{2}}^{n+1} - \overline{\rho u}_{i+\frac{1}{2}}^*) + \frac{\Delta t}{\Delta x} (\mathbf{R}^{\text{FV}} \cdot \bar{\mathbf{p}}_{i+1}^{n+\theta_{i+\frac{1}{2}}} - \mathbf{L}^{\text{FV}} \cdot \bar{\mathbf{p}}_i^{n+\theta_{i+\frac{1}{2}}}) = 0 \quad (5.38)$$

where  $\overline{\rho u}_{i+\frac{1}{2}}^*$  is the contribution of the explicit discretization of the nonlinear convective term; in addition, the tensors  $\mathbf{R}^{\text{FV}}$  and  $\mathbf{L}^{\text{FV}}$  (see Appendix A.3.2) approximate the derivative in  $x$ -direction. For the sake of clarity, we give the approximation of the pressure gradient  $\Delta x \partial p / \partial x$  with  $P = 2$

and  $\theta_{i \pm \frac{1}{2}} = 1$ :

$$\begin{aligned} \mathbf{R}^{\text{FV}} \cdot \bar{\mathbf{p}}_{i+\frac{1}{2}}^{n+1} - \mathbf{L}^{\text{FV}} \cdot \bar{\mathbf{p}}_{i-\frac{1}{2}}^{n+1} = \\ 5(\bar{p}_{i-\frac{1}{2},4}^{n+1} - \bar{p}_{i-\frac{1}{2},3}^{n+1} \bar{p}_{i-\frac{1}{2},5}^{n+1} - \bar{p}_{i-\frac{1}{2},4}^{n+1} \bar{p}_{i+\frac{1}{2},1}^{n+1} - \bar{p}_{i-\frac{1}{2},5}^{n+1} \bar{p}_{i+\frac{1}{2},2}^{n+1} - \bar{p}_{i+\frac{1}{2},1}^{n+1} \bar{p}_{i+\frac{1}{2},3}^{n+1} - \bar{p}_{i+\frac{1}{2},2}^{n+1} \bar{p}_{i+\frac{1}{2},4}^{n+1}). \end{aligned} \quad (5.39)$$

The semi-implicit sub-cell finite volume discretization of the total energy equation takes the form

$$\begin{aligned} \frac{1}{\gamma-1} \bar{\mathbf{p}}_i^{n+1,r+1} + \frac{\Delta t}{\Delta x} \left( \theta_{i+\frac{1}{2}} \mathbf{R}^{\text{FV}} \cdot \bar{\mathbf{h}}_{i+\frac{1}{2}}^{n+1,r} \bar{\rho} \mathbf{u}_{i+\frac{1}{2}}^{n+1,r+1} - \theta_{i-\frac{1}{2}} \mathbf{L}^{\text{FV}} \cdot \bar{\mathbf{h}}_{i-\frac{1}{2}}^{n+1,r} \bar{\rho} \mathbf{u}_{i-\frac{1}{2}}^{n+1,r+1} \right) \\ = \left( \bar{\rho} \mathbf{E}_i^* - \bar{\rho} \mathbf{k}_i^{n+1,r} \right) - \frac{\Delta t}{\Delta x} \left( (1 - \theta_{i+\frac{1}{2}}) \mathbf{R}^{\text{FV}} \cdot \bar{\mathbf{h}} \rho \mathbf{u}_{i+\frac{1}{2}}^n - (1 - \theta_{i-\frac{1}{2}}) \mathbf{L}^{\text{FV}} \cdot \bar{\mathbf{h}} \rho \mathbf{u}_{i-\frac{1}{2}}^n \right). \end{aligned} \quad (5.40)$$

Inserting eq. (5.38) into eq. (5.40) yields a linear three-diagonal system which at each Picard level reads

$$\mathcal{L}_i^{\text{FV}} \cdot \bar{\mathbf{p}}_{i-1}^{n+1,r+1} + \mathcal{C}_i^{\text{FV}} \cdot \bar{\mathbf{p}}_i^{n+1,r+1} + \mathcal{R}_i^{\text{FV}} \cdot \bar{\mathbf{p}}_{i+1}^{n+1,r+1} = \bar{\mathbf{b}}_i^r, \quad (5.41)$$

which can be solved either by a Thomas algorithm or by an iterative conjugate gradient method. In each Picard iteration the enthalpy, which is defined on the staggered mesh, has to be updated and one can use the following FV average in order to interpolate the pressure from the main grid to the dual mesh

$$\bar{\mathbf{p}}_{i+\frac{1}{2}} = (\mathbf{M}_L^{\text{FV}} \cdot \bar{\mathbf{p}}_i + \mathbf{M}_R^{\text{FV}} \cdot \bar{\mathbf{p}}_{i+1}). \quad (5.42)$$

Similarly, the kinetic energy is discretized on the main grid. So, in order to get the degrees of freedom of the velocity  $\bar{\mathbf{u}}_i$ , the projection from the dual grid to main one reads

$$\bar{\mathbf{u}}_i = (\mathbf{M}_L^{\text{FV}} \cdot \bar{\mathbf{u}}_{i-\frac{1}{2}} + \mathbf{M}_R^{\text{FV}} \cdot \bar{\mathbf{u}}_{i+\frac{1}{2}}). \quad (5.43)$$

For the matrices introduced in equations (5.42) and (5.43), see appendix A.3.2. Finally, the total energy update reads

$$\bar{\rho} \mathbf{E}_i^{n+1} = \bar{\rho} \mathbf{E}_i^* - \frac{\Delta t}{\Delta x} \cdot (\mathbf{R}^{\text{FV}} \cdot \bar{\mathbf{h}} \rho \mathbf{u}_{i+\frac{1}{2}}^{n+\theta} - \mathbf{L}^{\text{FV}} \cdot \bar{\mathbf{h}} \rho \mathbf{u}_{i-\frac{1}{2}}^{n+\theta}). \quad (5.44)$$

In this case, the CFL condition of the scheme is given by the stability condition for the computation of the explicit nonlinear convective terms:

$$\Delta t < \text{CFL} \frac{\Delta x_s}{\max(|u|)} \quad \text{with} \quad \text{CFL} < 1. \quad (5.45)$$

We remind that here we introduced an alternative data representation for the FV method which is necessary in order to combine the DG and FV schemes on staggered meshes for the *a posteriori* limiting strategy introduced in the next section. For the method discussed in this subsection, the solution algorithm is the same adopted for the 1D pure DG scheme in Section 5.1.2. For a better understanding of this particular semi-implicit FV method see [55, 90] or subsection 4.1.3 in Chapter 4.

### 5.1.4 MOOD algorithm and detection criteria - 1D case

In this method, we adopt the same limiting strategy adopted in Chapter 4 for staggered SIDG applied to the shallow water equations. Similarly to the *a posteriori* sub-cell limiter for explicit DG schemes exposed in [14, 59, 64], the starting point is the MOOD paradigm [39, 47, 48, 107]: at each time step the algorithm is composed of two stages. In the first one, the unlimited DG scheme presented in section 5.1.2 generates a so-called *candidate solution*  $\mathbf{Q}^{\circ, n+1} = (\rho^{\circ, n+1}, \rho u^{\circ, n+1}, \rho E^{\circ, n+1})^T$  at time  $t^{n+1}$ . Here, we use the circle symbol as apex for the candidate solution, in order to avoid confusion with the discretization of the nonlinear convective terms, which in this section are denoted by the star symbol. Then, the troubled zones are detected using physical and numerical admissibility criteria. In the cells flagged as troubled zone, the more robust staggered semi-implicit finite volume subcell scheme introduced in the previous section 5.1.3 is used, while we use the unlimited DG method in those control volumes which are not troubled. Afterwards, in the second part of the procedure, a valid solution at time  $t^n$  is recovered and then the update is carried out by a *mixed* scheme, which uses the more robust subcell finite volume method in troubled cells and the unlimited DG method in the other ones. To do this, the linear system for the pressure is reassembled and solved again, using the subcell FV method in troubled cells, and the unlimited DG scheme everywhere else. Moreover, if troubled cells have not been individuated after the first level of the algorithm, the scheme directly advances to the next time level.

#### 5.1.4.1 Data representation, projection and reconstruction

For a generic variable  $q$  of the approximated solution given by the DG scheme in the cell  $T_i$  at time  $t^n$ , the degrees of freedom used for the piecewise polynomial data representation are denoted by  $\hat{q}_i^n$ . Similarly  $\bar{q}_i^n$  denote the piecewise constant cell averages in the subcells  $T_{i,s}$ . These subcell averages  $\bar{q}_i^n$  are computed from  $\hat{q}_i^n$  using the following  $L_2$  projection

$$\bar{q}_{i,s}^n = \frac{1}{|T_{i,s}|} \int_{T_{i,s}} \phi_l(x) dx \hat{q}_{i,l}^n, \quad \forall T_{i,s} \in T_i, \quad (5.46)$$

and consequently we introduce the projection operator  $\mathcal{P}$  so that  $\bar{q}_i^n = \mathcal{P} \cdot \hat{q}_i^n$ . Then, one can gather back the piecewise constant sub-cell averages  $\bar{q}_{i,s}^n$  into the degrees of freedom  $\hat{q}_{i,l}^n$  of a high order DG polynomial by solving again (5.46). However, this procedure leads to an *over-determined* system because there are  $2P + 1$  equations for  $P + 1$  unknowns. In order to overcome this problem a *constrained least-squares* approach (see [57, 97]) is adopted. In particular, the linear constraint is integral conservation on the big cell  $T_i$  (see [64]), i.e.

$$\sum_j |T_{i,s}| \bar{q}_{i,s}^n = \int_{T_i} \phi_l(x) dx \hat{q}_{i,l}^n. \quad (5.47)$$

This operation is expressed in a tensorial formulation as  $\hat{q}_i^n = \mathcal{W} \cdot \bar{q}_i^n$ , with  $\mathcal{W} \cdot \mathcal{P} = \mathcal{I}$ , where  $\mathcal{I}$  denotes the identity matrix. For a schematic summary of these projection and reconstruction procedure, see Figure 5.3. Moreover, in some situations we will need to recover a polynomial from  $P + 1$  piecewise constant sub-cell data. Here we consider two situations: one is at the left part,  $T_i^L = \bigcup_{s=1}^{P+1} T_{i,s}$  (see Figs. 5.4), and the other at the right side,  $T_i^R = \bigcup_{s=P+1}^{2P+1} T_{i,s}$  (see Fig. 5.5). So we introduce the matrices  $\mathcal{W}_{\mathcal{L}}$  and  $\mathcal{W}_{\mathcal{R}}$  that are computed using reconstruction procedure

(5.46) only on the left and right part of cell  $T_i$ . Note that here the least squares procedure is not necessary because now the number of equations is equal to the number of unknowns.

#### 5.1.4.2 Detection criteria

In this subsection we discuss the physical and numerical admissibility criteria. All the quantities are analysed on the *main grid* and those variables which are located on the staggered dual grid are extrapolated onto the main grid. This operation is done by using the algorithm given in eq. (5.33). At first, we want that the candidate solution, given numerically by the unlimited staggered DG scheme, satisfies some *physical admissibility criteria* (PAD). For the Euler equations we check that both density and both pressure are non-negative, hence we require

$$\hat{\rho}_{i,l}^{\circ,n+1} \geq 0 \quad \text{and} \quad \hat{p}_{i,l}^{\circ,n+1} \geq 0. \quad (5.48)$$

In addition, similarly to [14, 59, 64], the detector is supplemented by *numerical admissibility criteria* (NAD) based on a relaxed version of the discrete maximum principle (DMP). For a generic variable  $q$  this criteria read as follows

$$\min_{\forall T_j \in \mathcal{V}_i} (\hat{q}_{j,l}^n) - \delta \leq \hat{q}_{i,l}^{\circ,n+1} \leq \max_{\forall T_j \in \mathcal{V}_i} (\hat{q}_{j,l}^n) + \delta, \quad (5.49)$$

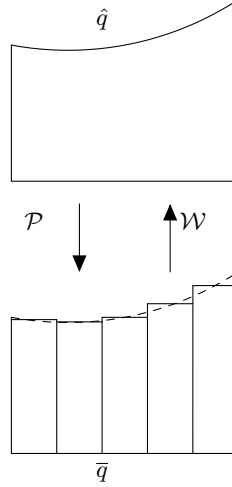
where  $\mathcal{V}_i = \{T_{i-1}, T_i, T_{i+1}\}$  is the set of neighbour cells of  $T_i$ , which is the set of Voronoi neighbours in the multidimensional case. In equation (5.49) we introduce two tolerances:  $\delta$  is a relaxation parameter that allows some very small oscillations or numerical noise; it reads

$$\delta = \max[\delta_0, \epsilon (\max_{\forall T_j \in \mathcal{V}_i} (\hat{q}_{j,l}^n) - \min_{\forall T_j \in \mathcal{V}_i} (\hat{q}_{j,l}^n))] \quad (5.50)$$

where  $\delta_0$  is chosen in the interval  $[10^{-4}, 10^{-3}]$  and typically  $\epsilon = 5 \cdot 10^{-4}$ . In this algorithm we check the DMP componentwise for all conserved variables. Moreover, an additional detection of troubled cells is carried out projecting the numerical quantities from the main grid to the dual mesh and then project them back onto the main grid. This procedure creates non-physical oscillations near discontinuities but it leads to a more robust detection of troubled cells, especially for the initial condition. We furthermore check the presence of NaN (not-a-number) values in the candidate solution, which are also a clear indicator for a cell to be troubled.

Successively, a cell is flagged as troubled if it does not respect the *a posteriori* admissibility criteria applied to the candidate solution  $\mathbf{Q}^{\circ,n+1}$ . We introduce an indicator parameter  $\beta_i$ . If  $\beta_i = 1$  is marked as troubled, the solution at time  $t^{n+1}$  will be recomputed using the finite volume scheme presented in Section 5.1.3. On the contrary, if  $\beta_i = 0$  in  $T_i$  it means that the cell is untroubled. In addition, if the cell  $T_i$  is flagged as troubled we also mark both dual control volumes  $T_{i-\frac{1}{2}}$  and  $T_{i+\frac{1}{2}}$  as troubled control volumes, see Figure 5.6. As consequence, we impose  $\beta_{i-\frac{1}{2}} = 1$  and  $\beta_{i+\frac{1}{2}} = 1$ .

Finally, for the sake of clarity we remind that at the first MOOD iteration the implicitness parameter  $\theta$  is a constant value  $\theta_{DG}$  in the whole domain. It is a user-defined parameter that is typically chosen very close to  $1/2$  in order to minimize the effects of the numerical viscosity caused by the time integration. At the second MOOD level, the value of the parameter  $\theta$  in the troubled cells of the dual grid is imposed equal to 1. Consequently, the new numerical solution is given by using a monotone and robust finite volume method in the troubled cells.



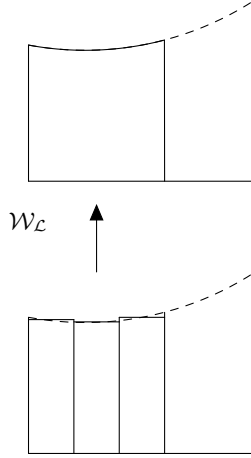
**Figure 5.3:** Projection and reconstruction operators - Operators  $\mathcal{P}$  and  $\mathcal{W}$

### 5.1.5 Sub-cell limiting of the semi-implicit DG scheme for the 1D Euler equations

Finally, we complete the derivation of the 1D algorithm for the finite volume limiting of the Euler equations. The computation of the nonlinear convective terms is carried out according to the MOOD strategy used for explicit high order schemes, see [14, 39, 47, 48, 59, 64, 107]. Hence, at the first MOOD step all the cells are flagged with  $\beta_i = 0$  and the first step of the algorithm is done using the pure unlimited DG algorithm explained in subsection 5.1.2.1. Later on, after the detection of the troubled control volumes, we use the explicit finite volume update given in 5.1.3 in order to recompute the solution in the cells marked with  $\beta = 1$  while for the explicit part of the algorithm in the non-troubled cells we can keep the results of the convective terms obtained in the first MOOD iteration.

Now, we consider the semi-implicit approximation of the one-dimensional total energy and momentum equations for the complete algorithm. For the general case we distinguish four situations. The first one is when the control volume  $T_i$  and the corresponding dual cells  $T_{i\pm 1/2}$  are not marked as troubled. In this case  $\beta_{i-\frac{1}{2}} = \beta_i = \beta_{i+\frac{1}{2}} = 0$  and the equations are discretized using the pure DG method introduced in subsection 5.1.2. The opposite case is when the troubled cell indicator  $\beta$  is equal to 1,  $\beta_{i-\frac{1}{2}} = \beta_i = \beta_{i+\frac{1}{2}} = 1$ , in all the control volumes  $T_{i-\frac{1}{2}}, T_i, T_{i+\frac{1}{2}}$ ; for this





**Figure 5.4:** Projection and reconstruction operators - Operator  $\mathcal{W}_{\mathcal{L}}$

situation, we use the finite volume method discussed in 5.1.3. Then we consider the two situations where the control volume  $T_i$  is not limited but in one of the corresponding dual control volumes  $T_{i-\frac{1}{2}}$  data are represented by finite volumes. In these cases the energy equation reads

$$\mathbf{M} \cdot \left( \frac{\widehat{\rho E}_i^{n+1} - \widehat{\rho E}_i^*}{\Delta t} \right) + \frac{\mathbf{R}_{\mathbf{e}}^{\text{DG}} \cdot \widehat{h}_{i+\frac{1}{2}}^{n+\theta} \widehat{\rho u}_{i+\frac{1}{2}}^{n+\theta} - \mathbf{L}_{\mathbf{e}}^{\text{Lim}} \cdot \overline{h \rho u}_{i-\frac{1}{2}}^{n-\theta}}{\Delta x} = 0 \quad (5.51)$$

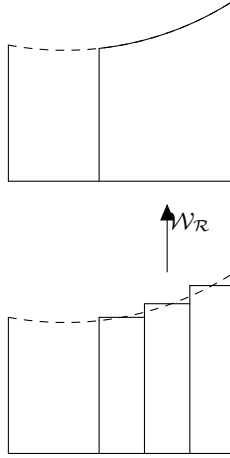
$$\text{with } \mathbf{L}_{\mathbf{e}}^{\text{Lim}} = \mathbf{L}_{\mathbf{e}}^{\text{DG}} \cdot \mathcal{W}_{\mathcal{R}}$$

for  $\beta_i = 0$   $\beta_{i+1} = 0$   $\beta_{i-1} = 1$  and

$$\mathbf{M} \cdot \left( \frac{\widehat{\rho E}_i^{n+1} - \widehat{\rho E}_i^*}{\Delta t} \right) + \frac{\mathbf{R}_{\mathbf{e}}^{\text{Lim}} \cdot \overline{h \rho u}_{i+\frac{1}{2}}^{n+\theta} - \mathbf{L}_{\mathbf{e}}^{\text{DG}} \cdot \widehat{h}_{i-\frac{1}{2}}^{n+\theta} \widehat{\rho u}_{i-\frac{1}{2}}^{n+\theta}}{\Delta x} = 0 \quad (5.52)$$

$$\text{with } \mathbf{R}_{\mathbf{e}}^{\text{Lim}} = \mathbf{R}_{\mathbf{e}}^{\text{DG}} \cdot \mathcal{W}_{\mathcal{L}}.$$

if  $\beta_i = 0$   $\beta_{i-1} = 0$   $\beta_{i+1} = 1$ . These two approximations are very similar to eq. (5.19), i.e. the discretization of the energy equation in the pure DG case. In fact, considering the scheme in Figure 5.7, the operator  $\mathcal{W}_{\mathcal{L}}$  reconstructs the polynomial data using the first  $P+1$  subcell averages of the



**Figure 5.5:** Projection and reconstruction operators - Operator  $\mathcal{W}_{\mathcal{R}}$

cell  $T_{i+\frac{1}{2}}$ . Similarly  $\mathcal{W}_{\mathcal{R}}$  reconstructs the polynomial data from the last  $P+1$  piecewise constant data of the control volume  $T_{i-\frac{1}{2}}$ , see Figure 5.52. In addition, in equations (5.51) and (5.51) we introduced matrices  $\mathbf{R}_{\mathbf{e}}^{\text{DG}}$  and  $\mathbf{L}_{\mathbf{e}}^{\text{DG}}$  that are universal tensors of rank 2, see Appendix A.3.3.

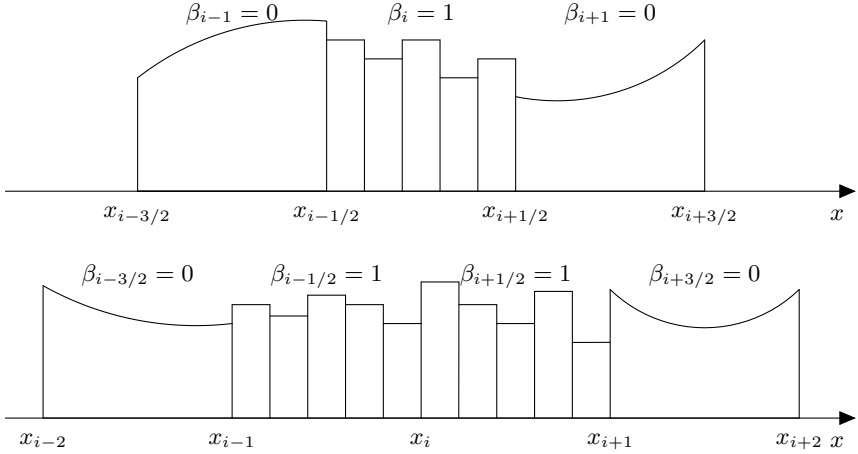
Now we discuss the two limited cases of the discrete momentum equation. Here, we consider the cases when  $T_{i+\frac{1}{2}}$  is a troubled cell and only one of the control volumes  $T_i$  and  $T_{i+1}$  is a limited control volume. Hence the numerical approximation of the momentum equation reads as follows

$$\overline{\rho \mathbf{u}}_{i+\frac{1}{2}}^{n+1} = \overline{\rho \mathbf{u}}_{i+\frac{1}{2}}^* - \frac{\Delta t}{\Delta x} (\mathbf{R}_{\mathbf{u}}^{\text{Lim}} \cdot \hat{\mathbf{p}}_{i+1}^{n+\theta_{i+\frac{1}{2}}} - \mathbf{L}^{\text{FV}} \cdot \hat{\mathbf{p}}_i^{n+\theta_{i+\frac{1}{2}}}) \quad \text{with} \quad \mathbf{R}_{\mathbf{u}}^{\text{Lim}} = \mathbf{R}^{\text{FV}} \cdot \mathcal{P} \quad (5.53)$$

when  $\beta_i = 0$  and  $\beta_{i+1} = 1$  (see Fig. 5.9)

$$\overline{\rho \mathbf{u}}_{i+\frac{1}{2}}^{n+1} = \overline{\rho \mathbf{u}}_{i+\frac{1}{2}}^* - \frac{\Delta t}{\Delta x} (\mathbf{R}^{\text{FV}} \cdot \hat{\mathbf{p}}_{i+1}^{n+\theta_{i+\frac{1}{2}}} - \mathbf{L}_{\mathbf{u}}^{\text{Lim}} \cdot \hat{\mathbf{p}}_i^{n+\theta_{i+\frac{1}{2}}}) \quad \text{with} \quad \mathbf{L}_{\mathbf{u}}^{\text{Lim}} = \mathbf{L}^{\text{FV}} \cdot \mathcal{P} \quad (5.54)$$

when  $\beta_i = 1$  and  $\beta_{i+1} = 0$  (see Fig. 5.10). Note that equation (5.53) and 5.53 can be seen as particular case of the discrete momentum equation (5.38) for the sub-cell finite volume



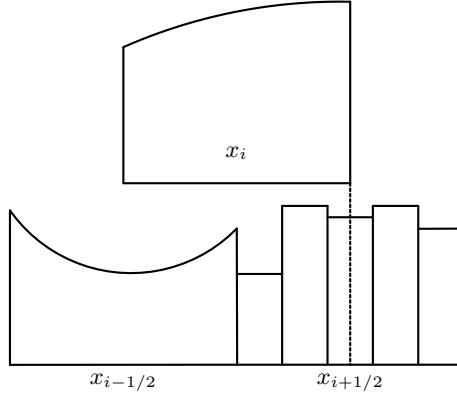
**Figure 5.6:** Staggered grids for the semi-implicit staggered DG scheme with sub-cell limiter active in cell  $T_i$ . Main grid used for the free surface (top) and staggered dual mesh for the velocity (bottom). If a cell  $T_i$  on the main grid is flagged as troubled, then also the two overlapping staggered velocity control volumes  $T_{i\pm\frac{1}{2}}$  are flagged as troubled. The data representation in troubled cells is changed from high order polynomials to piecewise constant subcell averages.

method. In addition, in eq. (5.53), the product  $\mathbf{R}_{\mathbf{u}}^{\text{Lim}} \cdot \tilde{\mathbf{p}}_{i+1}^{n+\theta_{i+\frac{1}{2}}} = \mathbf{R}^{\text{FV}} \cdot \tilde{\mathbf{p}}_{i+1}^{n+\theta_{i+\frac{1}{2}}}$  is equal to the multiplication of the tensor  $\mathbf{R}^{\text{FV}}$  with the vector of piecewise constant subcell averages  $\tilde{\mathbf{p}}_{i+1}^{n+\theta_{i+\frac{1}{2}}} = \mathcal{P} \cdot \tilde{\mathbf{p}}_{i+1}^{n+\theta_{i+\frac{1}{2}}}$  given by the projection of the DG degrees of freedom onto the set of the subcell finite volume degrees of freedom. This observation is valid also for the tensor product  $\mathbf{L}_{\mathbf{u}}^{\text{Lim}} \cdot \tilde{\mathbf{p}}_i^{n+\theta_{i+\frac{1}{2}}} = \mathbf{L}^{\text{FV}} \cdot \tilde{\mathbf{p}}_i^{n+\theta_{i+\frac{1}{2}}}$  in eq. (5.54).

The substitution of the momentum equation into the energy equation yields the three-diagonal linear system in eq. (5.55).

$$\mathcal{L}_i^{\text{Lim}} \cdot \tilde{\mathbf{p}}_{i-1}^{n+1, r+1} + \mathcal{C}_i^{\text{Lim}} \cdot \tilde{\mathbf{p}}_i^{n+1, r+1} + \mathcal{R}_i^{\text{Lim}} \cdot \tilde{\mathbf{p}}_{i+1}^{n+1, r+1} = \tilde{\mathbf{b}}_i^r. \quad (5.55)$$

We denote with the symbol  $\tilde{\mathbf{p}}_i^{n+1}$  the generic degrees of freedom of the pressure which represent either a DG polynomial or a set of subcell averages depending on the value of the troubled cell



**Figure 5.7:** Control volumes for the energy equation for the staggered semi-implicit DG scheme in the limited case using  $P = 2$ , i.e.  $N_s = 2P + 1 = 5$ . DG method used in  $T_i$  and  $T_{i-\frac{1}{2}}$ ; finite volumes in  $T_{i+\frac{1}{2}}$ .

indicator  $\beta$ .

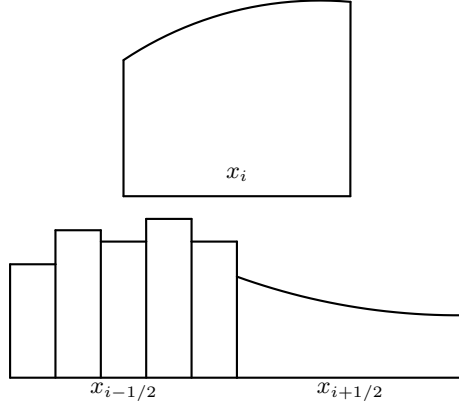
In the following, we will give the expressions for the blocks  $\mathcal{L}_i^{\text{Lim}}$ ,  $\mathcal{C}_i^{\text{Lim}}$ ,  $\mathcal{R}_i^{\text{Lim}}$  and for the right hand side term  $\tilde{\mathbf{b}}_i$ . For clarity,  $\mathcal{C}_i^{\text{Lim}}$  is divided in two contributions

$$\mathcal{C}_i^{\text{Lim}} = \mathcal{C}_i^{0,\text{Lim}} + \mathcal{C}_i^{x,\text{Lim}}, \quad (5.56)$$

where  $\mathcal{C}_i^{0,\text{Lim}}$  is equal to the mass matrix in the DG case or equal to the identity matrix in the FV situation:

$$\mathcal{C}_i^{0,\text{Lim}} = \begin{cases} \mathbf{M} & \text{if } \beta_i = 0, \\ \mathbf{I} & \text{if } \beta_i = 1. \end{cases} \quad (5.57)$$

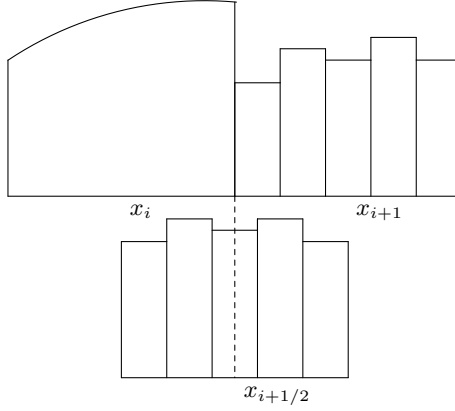
The remaining blocks  $\mathcal{L}_i^{\text{Lim}}$ ,  $\mathcal{C}_i^{x,\text{Lim}}$ ,  $\mathcal{R}_i^{\text{Lim}}$  can be expressed considering four different cases de-



**Figure 5.8:** Control volumes for the energy equation for the staggered semi-implicit DG scheme in the limited case using  $P = 2$ , i.e.  $N_s = 2P + 1 = 5$ . DG method used in  $T_i$  and  $T_{i+\frac{1}{2}}$ ; finite volumes in  $T_{i-\frac{1}{2}}$ .

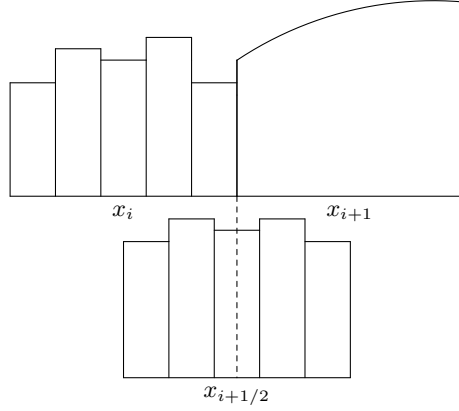
pending on the value of  $\beta$

$$\mathcal{L}_i^{\text{Lim}} = -\frac{\Delta t^2}{\Delta x^2} \theta_{i-\frac{1}{2}}^2 \begin{cases} \mathbf{L}_e^{\text{DG}} \cdot \hat{\mathbf{h}}_{\sim_{i-\frac{1}{2}}}^{n+1} \cdot \mathbf{M}^{-1} \cdot \mathbf{L}_u^{\text{DG}} & \text{if } \beta_{i-1} = \beta_i = 0, \\ \mathbf{L}^{\text{FV}} \cdot \text{diag}(\hat{\mathbf{h}}_{\sim_{i-\frac{1}{2}}}^{n+1}) \cdot \mathbf{L}^{\text{FV}} & \text{if } \beta_{i-1} = \beta_i = 1, \\ \mathbf{L}_e^{\text{Lim}} \cdot \text{diag}(\hat{\mathbf{h}}_{\sim_{i-\frac{1}{2}}}^{n+1}) \cdot \mathbf{L}^{\text{FV}} & \text{if } \beta_{i-1} = 1, \beta_i = 0, \\ \mathbf{L}^{\text{FV}} \cdot \text{diag}(\hat{\mathbf{h}}_{\sim_{i-\frac{1}{2}}}^{n+1}) \cdot \mathbf{L}_u^{\text{Lim}} & \text{if } \beta_{i-1} = 0, \beta_i = 1, \end{cases} \quad (5.58)$$



**Figure 5.9:** Control volumes for the momentum equation for the staggered semi-implicit DG scheme in the limited case using  $P = 2$  ( $N_s = 5$ ). DG in  $T_i$ ; FV in  $T_{i+1}$  and in  $T_{i+\frac{1}{2}}$

$$C_i^{x,\text{Lim}} = \frac{\Delta t^2}{\Delta x^2} \left\{ \begin{array}{l} \theta_{i-\frac{1}{2}}^2 \mathbf{R}_e^{\text{DG}} \cdot \hat{\mathbf{h}}_{\sim i-\frac{1}{2}}^{n+1} \cdot \mathbf{M}^{-1} \cdot \mathbf{L}_u^{\text{DG}} + \theta_{i+\frac{1}{2}}^2 \mathbf{L}_e^{\text{DG}} \cdot \hat{\mathbf{h}}_{\sim i+\frac{1}{2}}^{n+1} \cdot \mathbf{M}^{-1} \cdot \mathbf{R}_u^{\text{DG}} \\ \text{if } \beta_i = \beta_{i-\frac{1}{2}} = \beta_{i+\frac{1}{2}} = 0, \\ \theta_{i-\frac{1}{2}}^2 \mathbf{L}_e^{\text{FV}} \cdot \text{diag}(\overline{\mathbf{h}}_{\sim i-\frac{1}{2}}^{n+1}) \cdot \mathbf{R}^{\text{FV}} + \theta_{i+\frac{1}{2}}^2 \mathbf{R}^{\text{FV}} \cdot \text{diag}(\overline{\mathbf{h}}_{\sim i+\frac{1}{2}}^{n+1}) \cdot \mathbf{L}^{\text{FV}} \\ \text{if } \beta_i = \beta_{i-\frac{1}{2}} = \beta_{i+\frac{1}{2}} = 1, \\ \theta_{i-\frac{1}{2}}^2 \mathbf{R}_e^{\text{DG}} \cdot \hat{\mathbf{h}}_{\sim i-\frac{1}{2}}^{n+1} \cdot \mathbf{M}^{-1} \cdot \mathbf{L}_u^{\text{DG}} + \theta_{i+\frac{1}{2}}^2 \mathbf{R}_e^{\text{Lim}} \cdot \text{diag}(\overline{\mathbf{h}}_{\sim i+\frac{1}{2}}^{n+1}) \cdot \mathbf{L}_u^{\text{Lim}} \\ \text{if } \beta_i = \beta_{i-\frac{1}{2}} = 0, \beta_{i+\frac{1}{2}} = 1, \\ \theta_{i-\frac{1}{2}}^2 \mathbf{L}_e^{\text{Lim}} \cdot \text{diag}(\overline{\mathbf{h}}_{\sim i-\frac{1}{2}}^{n+1}) \cdot \mathbf{R}_u^{\text{Lim}} + \theta_{i+\frac{1}{2}}^2 \mathbf{L}_e^{\text{DG}} \cdot \hat{\mathbf{h}}_{\sim i+\frac{1}{2}}^{n+1} \cdot \mathbf{M}^{-1} \cdot \mathbf{R}_u^{\text{DG}} \\ \text{if } \beta_i = \beta_{i+\frac{1}{2}} = 0, \beta_{i-\frac{1}{2}} = 1, \end{array} \right. \quad (5.59)$$



**Figure 5.10:** Control volumes for the momentum equation for the staggered semi-implicit DG scheme in the limited case using  $P = 2$  ( $N_s = 5$ ). DG in  $T_{i+1}$ ; FV in  $T_i$  and in  $T_{i+\frac{1}{2}}$

$$\mathcal{R}_i^{\text{Lim}} = -\frac{\Delta t^2}{\Delta x^2} \theta_{i+\frac{1}{2}}^2 \begin{cases} \mathbf{R}_e^{\text{DG}} \cdot \tilde{\mathbf{h}}_{i+\frac{1}{2}}^{n+1} \cdot \mathbf{M}^{-1} \cdot \mathbf{R}_u^{\text{DG}} & \text{if } \beta_{i+1} = \beta_i = 0, \\ \mathbf{R}^{\text{FV}} \cdot \text{diag}(\tilde{\mathbf{h}}_{i+\frac{1}{2}}^{n+1}) \cdot \mathbf{R}^{\text{FV}} & \text{if } \beta_{i+1} = \beta_i = 1, \\ \mathbf{R}_e^{\text{Lim}} \cdot \text{diag}(\tilde{\mathbf{h}}_{i+\frac{1}{2}}^{n+1}) \cdot \mathbf{R}^{\text{FV}} & \text{if } \beta_{i+1} = 1, \beta_i = 0, \\ \mathbf{R}^{\text{FV}} \cdot \text{diag}(\tilde{\mathbf{h}}_{i+\frac{1}{2}}^{n+1}) \cdot \mathbf{R}_u^{\text{Lim}} & \text{if } \beta_{i+1} = 0, \beta_i = 1. \end{cases} \quad (5.60)$$

Finally the right hand side term reads

$$\tilde{\mathbf{b}}_i^n = \tilde{\mathbf{b}}_i^{0,n} + \tilde{\mathbf{b}}_i^{x,n} \quad (5.61)$$

with

$$\tilde{\mathbf{b}}_i^{0,n} = \begin{cases} \mathbf{M} \cdot (\widehat{\rho \mathbf{E}}_i^* - \widehat{\rho \mathbf{k}}_i^{n+1}) & \text{if } \beta_i = 0, \\ \widehat{\rho \mathbf{E}}_i^* - \widehat{\rho \mathbf{k}}_i^{n+1} & \text{if } \beta_i = 1, \end{cases} \quad (5.62)$$

and with

$$\tilde{b}_i^{x,n} = -\frac{\Delta t}{\Delta x} \left\{ \begin{array}{l} (\theta_{i+\frac{1}{2}} \mathbf{R}_e^{\text{DG}} \cdot \tilde{\mathbf{h}}_{i+\frac{1}{2}}^{n+1} \hat{\mathbf{G}}_{i+\frac{1}{2}}^n - \theta_{i-\frac{1}{2}} \mathbf{L}_e^{\text{DG}} \cdot \tilde{\mathbf{h}}_{i-\frac{1}{2}}^{n+1} \hat{\mathbf{G}}_{i-\frac{1}{2}}^n) \\ + ((1 - \theta_{i+\frac{1}{2}}) \mathbf{R}_e^{\text{DG}} \cdot \hat{\mathbf{h}}_{i+\frac{1}{2}}^n \widehat{\rho \mathbf{u}}_{i+\frac{1}{2}}^n - (1 - \theta_{i-\frac{1}{2}}) \mathbf{L}_e^{\text{DG}} \cdot \hat{\mathbf{h}}_{i-\frac{1}{2}}^n \widehat{\rho \mathbf{u}}_{i-\frac{1}{2}}^n) \\ \text{if } \beta_i = \beta_{i-\frac{1}{2}} = \beta_{i+\frac{1}{2}} = 0, \\ (\theta_{i+\frac{1}{2}} \mathbf{R}^{\text{FV}} \cdot \tilde{\mathbf{h}}_{i+\frac{1}{2}}^n \overline{\mathbf{G}}_{i+\frac{1}{2}}^n - \theta_{i-\frac{1}{2}} \mathbf{L}^{\text{FV}} \cdot \tilde{\mathbf{h}}_{i-\frac{1}{2}}^n \overline{\mathbf{G}}_{i-\frac{1}{2}}^n) \\ + ((1 - \theta_{i+\frac{1}{2}}) \mathbf{R}^{\text{FV}} \cdot \overline{\mathbf{h}} \rho \mathbf{u}_{i+\frac{1}{2}}^n - (1 - \theta_{i-\frac{1}{2}}) \mathbf{L}^{\text{FV}} \cdot \overline{\mathbf{h}} \rho \mathbf{u}_{i-\frac{1}{2}}^n) \\ \text{if } \beta_i = \beta_{i-\frac{1}{2}} = \beta_{i+\frac{1}{2}} = 1, \\ (\theta_{i+\frac{1}{2}} \mathbf{R}_e^{\text{Lim}} \cdot \tilde{\mathbf{h}}_{i+\frac{1}{2}}^{n+1} \overline{\mathbf{G}}_{i+\frac{1}{2}}^n - \theta_{i-\frac{1}{2}} \mathbf{L}_e^{\text{DG}} \cdot \tilde{\mathbf{h}}_{i-\frac{1}{2}}^{n+1} \hat{\mathbf{G}}_{i-\frac{1}{2}}^n) \\ + ((1 - \theta_{i+\frac{1}{2}}) \mathbf{R}_e^{\text{Lim}} \cdot \overline{\mathbf{h}} \rho \mathbf{u}_{i+\frac{1}{2}}^n - (1 - \theta_{i-\frac{1}{2}}) \mathbf{L}_e^{\text{DG}} \cdot \hat{\mathbf{h}}_{i-\frac{1}{2}}^n \widehat{\rho \mathbf{u}}_{i-\frac{1}{2}}^n) \\ \text{if } \beta_i = \beta_{i-\frac{1}{2}} = 0, \beta_{i+\frac{1}{2}} = 1, \\ (\theta_{i+\frac{1}{2}} \mathbf{R}_e^{\text{DG}} \cdot \tilde{\mathbf{h}}_{i+\frac{1}{2}}^{n+1} \hat{\mathbf{G}}_{i+\frac{1}{2}}^n - \theta_{i-\frac{1}{2}} \mathbf{L}_e^{\text{Lim}} \cdot \tilde{\mathbf{h}}_{i-\frac{1}{2}}^{n+1} \overline{\mathbf{G}}_{i-\frac{1}{2}}^n) \\ + ((1 - \theta_{i+\frac{1}{2}}) \mathbf{R}_e^{\text{DG}} \cdot \hat{\mathbf{h}}_{i+\frac{1}{2}}^n \widehat{\rho \mathbf{u}}_{i+\frac{1}{2}}^n - (1 - \theta_{i-\frac{1}{2}}) \mathbf{L}_e^{\text{Lim}} \cdot \overline{\mathbf{h}} \rho \mathbf{u}_{i-\frac{1}{2}}^n) \\ \text{if } \beta_i = \beta_{i+\frac{1}{2}} = 0, \beta_{i-\frac{1}{2}} = 1. \end{array} \right. \quad (5.63)$$

We conclude with the algorithms for the interpolation from the main grid to the dual mesh

$$\tilde{\rho}_{i+\frac{1}{2}} = \left\{ \begin{array}{ll} \mathbf{M}^{-1} \cdot (\mathbf{M}_L^{\text{DG}} \cdot \hat{\rho}_i + \mathbf{M}_R^{\text{DG}} \cdot \hat{\rho}_{i+1}) & \text{if } \beta_{i+\frac{1}{2}} = \beta_i = \beta_{i+1} = 0, \\ (\mathbf{M}_L^{\text{FV}} \cdot \bar{\rho}_i + \mathbf{M}_R^{\text{FV}} \cdot \bar{\rho}_{i+1}) & \text{if } \beta_{i+\frac{1}{2}} = \beta_i = \beta_{i+1} = 1, \\ (\mathbf{M}_L^{\text{FV}} \cdot \mathcal{P} \cdot \hat{\rho}_i + \mathbf{M}_R^{\text{FV}} \cdot \bar{\rho}_{i+1}) & \text{if } \beta_{i+\frac{1}{2}} = \beta_{i+1} = 1, \beta_i = 0, \\ (\mathbf{M}_L^{\text{FV}} \cdot \bar{\rho}_i + \mathbf{M}_R^{\text{FV}} \cdot \mathcal{P} \cdot \hat{\rho}_{i+1}) & \text{if } \beta_{i+\frac{1}{2}} = \beta_i = 1, \beta_{i+1} = 0 \end{array} \right. \quad (5.64)$$

and with the one for the opposite projection from the staggered grid to main one

$$\tilde{\mathbf{U}}_i = \left\{ \begin{array}{ll} \mathbf{M}^{-1} \cdot (\mathbf{M}_L^{\text{DG}} \cdot \hat{\mathbf{U}}_{i-\frac{1}{2}} + \mathbf{M}_R^{\text{DG}} \cdot \hat{\mathbf{U}}_{i+\frac{1}{2}}) & \text{if } \beta_i = \beta_{i-\frac{1}{2}} = \beta_{i+\frac{1}{2}} = 0, \\ (\mathbf{M}_L^{\text{FV}} \cdot \bar{\mathbf{U}}_{i-\frac{1}{2}} + \mathbf{M}_R^{\text{FV}} \cdot \bar{\mathbf{U}}_{i+\frac{1}{2}}) & \text{if } \beta_i = \beta_{i-\frac{1}{2}} = \beta_{i+\frac{1}{2}} = 1, \\ \mathbf{M}^{-1} \cdot (\mathbf{M}_L^{\text{DG}} \cdot \hat{\mathbf{U}}_{i-\frac{1}{2}} + \mathbf{M}_R^{\text{DG}} \cdot \mathcal{W}_L \cdot \bar{\mathbf{U}}_{i+\frac{1}{2}}) & \text{if } \beta_i = \beta_{i-\frac{1}{2}} = 0, \beta_{i+\frac{1}{2}} = 1, \\ \mathbf{M}^{-1} \cdot (\mathbf{M}_L^{\text{DG}} \cdot \mathcal{W}_R \cdot \bar{\mathbf{U}}_{i-\frac{1}{2}} + \mathbf{M}_R^{\text{DG}} \cdot \hat{\mathbf{U}}_{i+\frac{1}{2}}) & \text{if } \beta_i = \beta_{i+\frac{1}{2}} = 0, \beta_{i-\frac{1}{2}} = 1, \end{array} \right. \quad (5.65)$$

Note that, due to the choice of  $2P + 1$  subcell averages, both for the explicit DG scheme and both for the explicit subgrid finite volume scheme the CFL stability condition gives the same maximum admissible time step  $\Delta t$ . The extension of the present algorithm to two space dimensions is quite straightforward and follows the ideas outlined in Chapter 4 for the shallow water equations. Therefore, further details about the 2D case can be omitted here.



## 5.2 Numerical tests for the one-dimensional model

In this section we carry out numerical test cases in order to validate the numerical method presented in section 5.1. The aim of the first test, which is reported in subsection 5.2.1, is to verify the correctness of the tensors for the pure DG scheme. Then, in subsection 5.2.2 the numerical results are compared against the analytical solution of the Riemann Problem in order to check the robustness of the limiter in presence of discontinuities. We remind that in all the benchmarks we do not use neither artificial viscosity nor smoothing.

### 5.2.1 Advection of a smooth density wave in 1D

Here, we consider a test in order to check the correct formulation of the pure DG scheme derived in subsection 5.1.2. As initial condition, in the computational domain  $\Omega_x = [-1, +1]$  we impose a smooth Gaussian profile for the density

$$\rho(x, 0) = \rho_0(1 + e^{-\frac{0.5x^2}{0.1^2}}) \quad (5.66)$$

with  $\rho_0 = 0.01$  and a constant value for the velocity  $U(x, 0) = U_0 = 2$  and for the pressure  $p(x, 0) = p_0 = 1$ . Consequently, the density wave is transported and, if we prescribe periodic boundary conditions, the exact solution at the final time  $t_{end} = 1$  coincides with the initial condition. The computational domain is  $\Omega_x = [-1, 1]$  and we carry out two simulations: in the first the grid has 50 cells with a polynomial degree  $P = 5$  while in the second case there are 300 control volumes with  $P = 0$  that corresponds to a low order finite volume scheme. Note that in both cases the degrees of freedom are equal to 300. The implicitness parameter  $\theta$  is fixed equal to 0.55. The results are presented in Figure 5.11. It is possible to observe, that for this smooth test the results obtained using the high order semi-implicit DG scheme are very accurate and they have an excellent agreement with the analytical solution. On the contrary, for the finite volume case we notice that the numerical solution for density is very diffused. In addition, in the plots for the velocity and for the pressure, the numerical noise of the FV method is several orders of magnitude bigger than the noise produced by the DG scheme with  $P = 5$ .

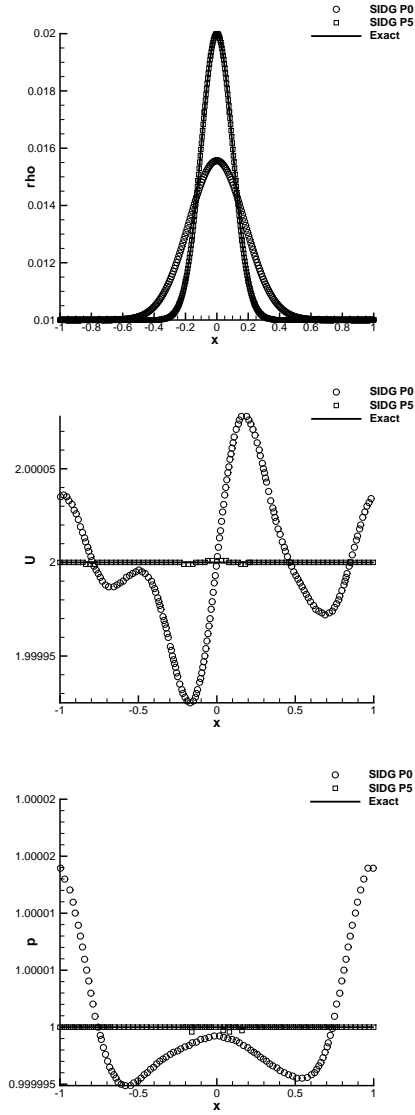
### 5.2.2 One-dimensional Riemann problems

Here we apply the novel semi-implicit DG scheme to a set of Riemann problems taken from [55, 140, 148]. These simulations are carried out in order to check the correct propagation of shock waves and to verify that the new method does not produce spurious or dispersive oscillations in the vicinity of discontinuities.

The computational domain is  $\Omega_x = [-0.5, +0.5]$  and the initial condition consists in a discontinuity centred in  $x_0$ ,

$$\mathbf{V}(x, 0) = \begin{cases} \mathbf{V}_L & \text{if } x \leq x_0 \\ \mathbf{V}_R & \text{if } x > x_0, \end{cases} \quad (5.67)$$

where  $\mathbf{V}$  indicates the vector of the primitive variables  $\mathbf{V} = [\rho, u, p]$ . In all the simulations  $N_x$  is equal to 100 and consequently  $\Delta x = 0.01$ . In addition the simulation run up to  $t = t_{end}$ . The parameters for the initial condition are listed in Table 5.1 and the results are shown in Figs. 5.12 - 5.18. The exact solution is given by the exact Riemann solver that can be found in the well-known textbook of Toro [148].



**Figure 5.11:** Reference solutions and numerical solutions for the problem at  $t = 1$ .

Case	$\rho_L$	$U_L$	$p_L$	$\rho_R$	$U_R$	$p_R$	$x_0$	$t_{end}$
RP1	1.0	0.0	1.0	0.125	0.0	0.1	0.0	0.2
RP2	0.445	0.698	3.528	0.5	0.0	0.571	0.0	0.14
RP3	1.0	0.0	1000.0	1.0	0.0	0.01	0.1	0.012
RP4	5.9	19.6	460.9	5.9	-6.2	46.1	0.0	0.035
RP5	1.0	-1.0	0.4	1.0	+1.0	0.4	0.0	0.15
RP6	1.0	+2.0	0.1	1.0	-2.0	0.1	0.0	0.5
RP7	1.0	0.75	1.0	0.125	0.0	0.1	-0.1	0.2

**Table 5.1:** Riemann problems for semi-implicit DG on staggered grid with sub-cell limiter - Left and right states for  $\rho, U$  and  $p$ ,  $x_0$ , number of control volume  $N_x$ ,  $\Delta x$  and final time of the simulations  $t_{end}$ .

RP1 (see Fig. 5.12) is known as Sod shock tube [131]. We observe an excellent agreement with the reference solution and, moreover, the limiter is activated only in the region of the shock wave, while the rarefaction and the contact waves are solved by using the pure unlimited DG method.

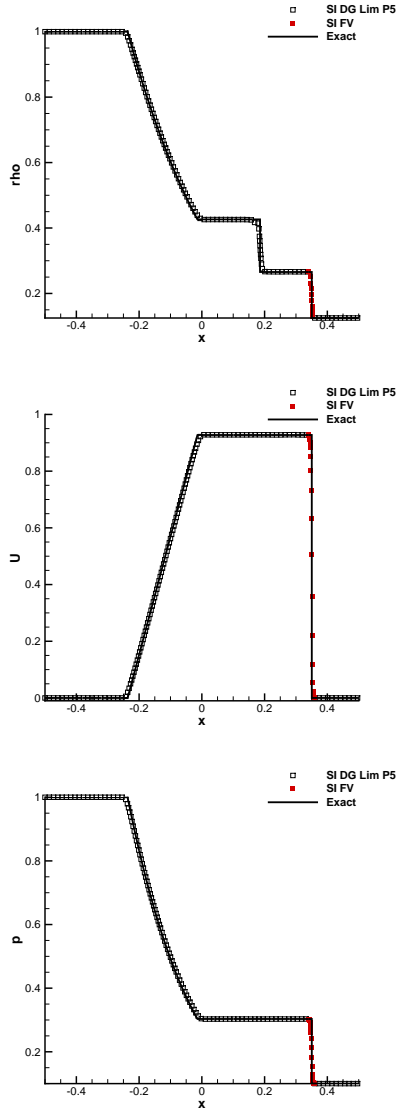
RP2 (see Fig. 5.13) is called Lax problem [103]. Also here we see a very good fitting with the exact solution of the Riemann problem. The new semi-implicit DG method with sub-cell limiter behaves very well and the limiter works only on the shock waves.

RP3 and RP4 (see Figures 5.14 and 5.15) are taken from [148]; these test cases involve very strong shock waves and consequently they are suitable for checking the robustness of the schemes. In RP3 we observe a very small overshoot on the shock wave however in the other parts of the domain the fitting is very good. Also in RP4 (see Fig. 5.15) the shocks are very well resolved, but we observe some additional troubled cells, probably because of some numerical noise in the plateau region.

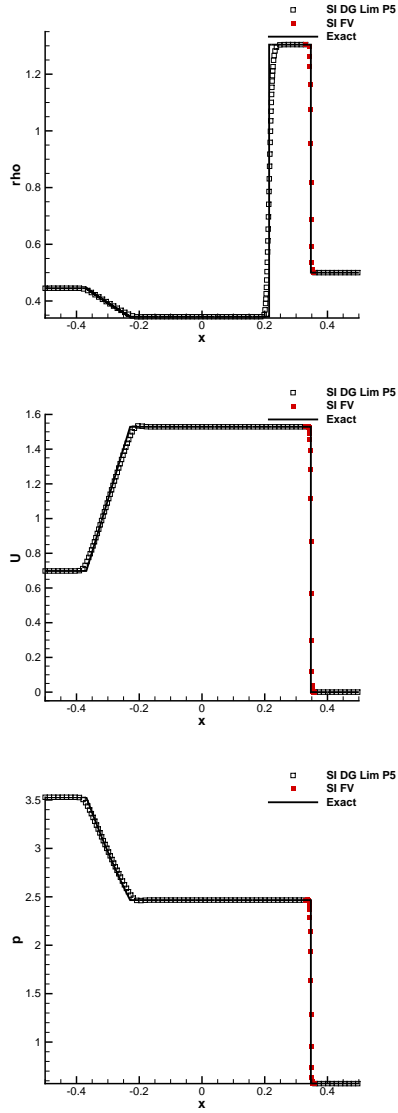
RP5 (see Fig. 5.16) consists of two symmetric rarefactions waves. We notice just a small wiggle close to the origin, but in the other zones the agreement is very good. Moreover, troubled cells are detected only in the first part of the simulation because the initial condition is discontinuous.

In RP6 (see Fig. 5.17) two jets collide and a double shock is generated; we observe that the waves are perfectly symmetric and the troubled cells are optimally detached and solved without any Gibbs phenomena. The glitch in the density has been observed also in [55]

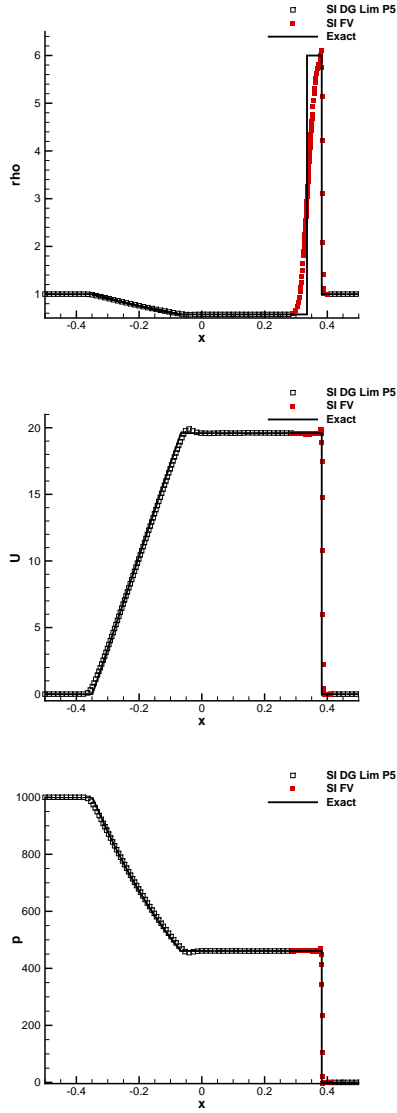
RP7 (see Fig. 5.18) is a modified Sod problem proposed by Toro in [148] in order to verify the presence of an entropy glitch inside the left moving rarefaction, which is evident in some entropy-violating approximate Riemann solvers, such as the one of Roe [125]. Again, the agreement is very good and we can conclude that the finite volume sub-cell limiter applied to semi-implicit DG is a very good shock capturing strategy. Moreover, for all these test cases, numerical experiments have shown that if the limiter is not activated, the simulation stops and it does not arrive at the final time  $t_{end}$ . We attribute this fact to the numerical viscosity of the pure DG method, which so low that it is not able to smooth out the Gibbs oscillations and the code terminates because NaN values were produced after the generation of negative densities and pressures.



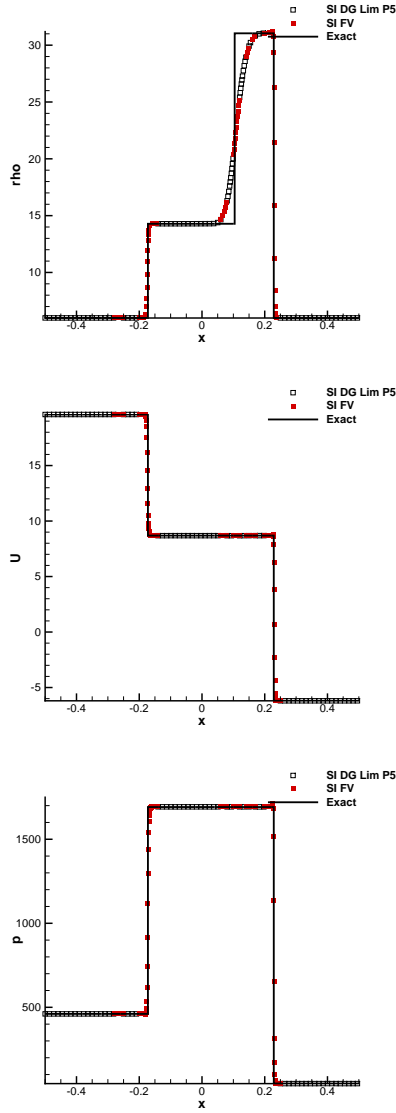
**Figure 5.12:** RP1 - Reference solutions and numerical solutions for  $\rho$ ,  $U$  and  $p$  at  $t_{end}$



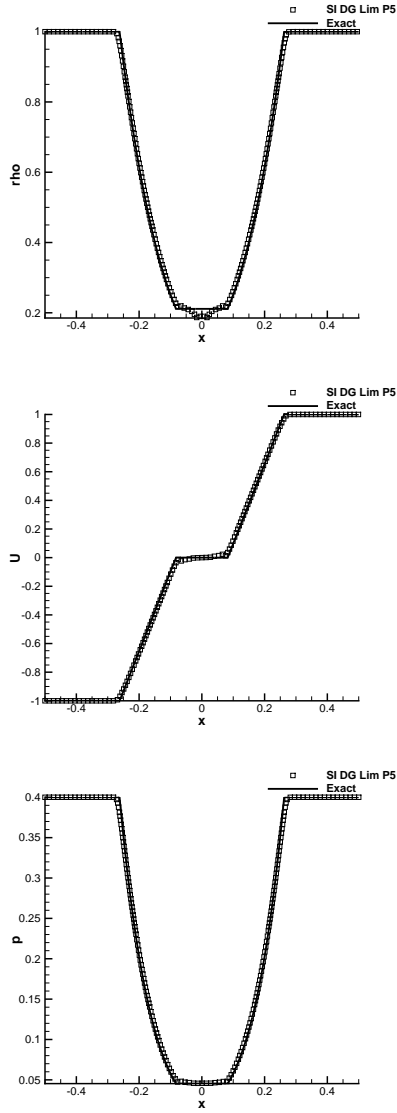
**Figure 5.13:** RP2 - Reference solutions and numerical solutions for  $\rho$ ,  $U$  and  $p$  at  $t_{end}$



**Figure 5.14:** RP3 - Reference solutions and numerical solutions for  $\rho$ ,  $U$  and  $p$  at  $t_{end}$



**Figure 5.15:** RP4 - Reference solutions and numerical solutions for  $\rho$ ,  $U$  and  $p$  at  $t_{end}$



**Figure 5.16:** RP5 - Reference solutions and numerical solutions for  $\rho$ ,  $U$  and  $p$  at  $t_{end}$



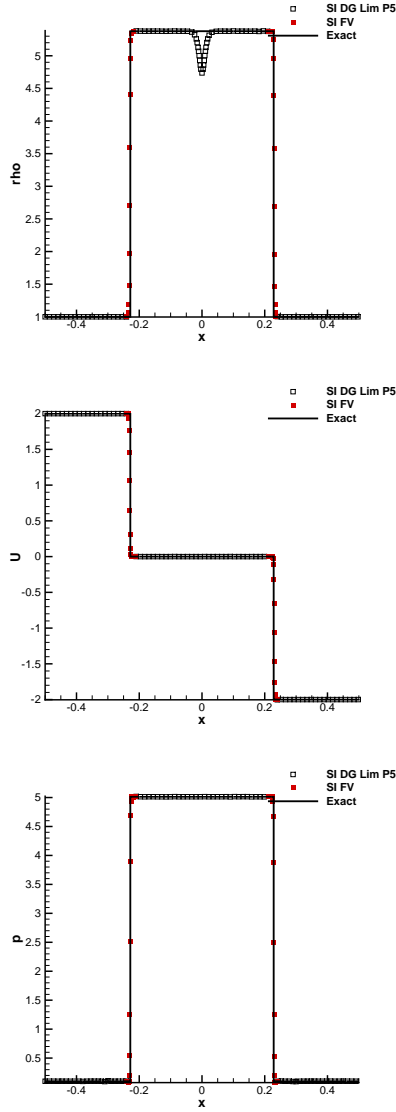
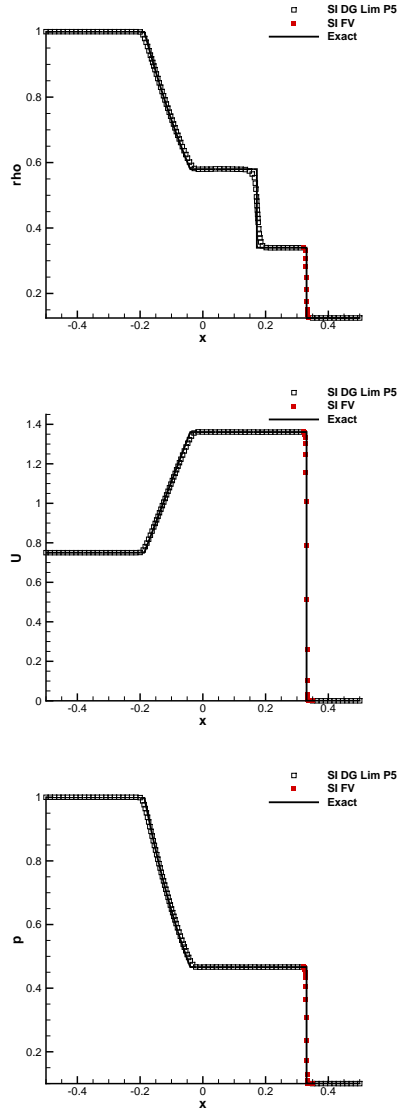


Figure 5.17: RP6 - Reference solutions and numerical solutions for  $\rho$ ,  $U$  and  $p$  at  $t_{end}$



**Figure 5.18:** RP7 - Reference solutions and numerical solutions for  $\rho$ ,  $U$  and  $p$  at  $t_{end}$

## 5.3 Numerical tests for the two-dimensional model

Finally, we carry out the numerical validation of the two-dimensional semi-implicit staggered DG scheme with *a posteriori* subcell finite volume limiting. At first we consider two smooth problems. In particular in subsection 5.3.1 we compute a numerical convergence table of the 2D scheme; then the test case in 5.3.2 can be seen as a sanity check, analogous to the one of subsection 5.2.1, while in 5.3.3 we perform a benchmark considering the low Mach number regime. We conclude with some 2D Riemann problems in sub-section 5.3.5, where the use of the limiter is fundamental in order to obtain a clean and accurate solution of the PDEs.

### 5.3.1 Isentropic vortex

The first two-dimensional test that we consider is the well known *isentropic vortex* (see [130]). For the 2D Euler equations, this is a fundamental benchmark because one can check the correctness of the implementation and compute numerical convergence rates. In addition, shocks and discontinuities are not involved and the boundary conditions are simply periodic. The initial condition for the density and for the pressure are  $\rho(x, y, 0) = \rho_\infty + \delta\rho$  and  $p(x, y, 0) = p_\infty + \delta p$ , where  $\delta\rho$  and  $\delta p$  read as follows

$$\delta\rho = (1 + \delta T)^{\frac{1}{\gamma-1}} - 1 \quad \delta p = (1 + \delta T)^{\frac{\gamma}{\gamma-1}} - 1 \quad \text{with} \quad \delta T = \frac{(\gamma-1)\beta^2}{8\gamma\pi^2} e^{1-r^2} \quad (5.68)$$

where we introduced the radial coordinate  $r = \sqrt{(x-x_0)^2 + (y-y_0)^2}$  and  $\beta$  is the vortex strength, here imposed equal to 5. The initial velocity is  $[U(x, y, 0), V(x, y, 0)] = [U_\infty + \delta u, V_\infty + \delta v]$  where  $\delta u$  and  $\delta v$  are given as

$$\delta u = -\frac{\beta}{2\pi} e^{\frac{1-r^2}{2}} (y - y_0) \quad \delta v = \frac{\beta}{2\pi} e^{\frac{1-r^2}{2}} (x - x_0). \quad (5.69)$$

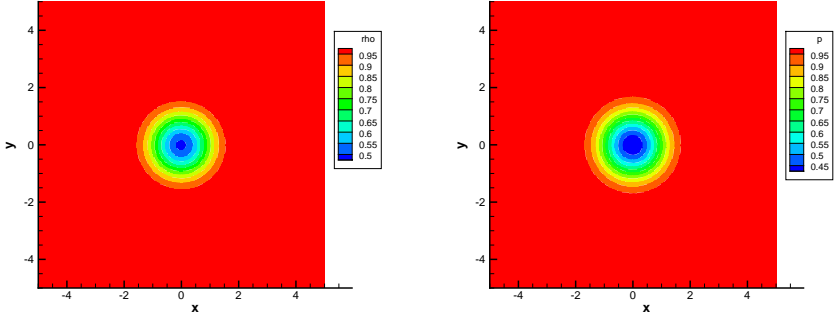
If the undisturbed velocities  $U_\infty$  and  $V_\infty$  are imposed equal to 0 the vortex becomes stationary. For our test, we therefore set  $U_\infty = V_\infty = 0$ . In this particular case we compute the  $L_2$  error between the numerical and the exact solutions on a series of successively refined grids, characterized by square cells in a domain  $\Omega_{xy} = [-5, +5] \times [-5, +5]$  with  $N_x = N_y$  and  $x_0 = y_0 = 0$ . For this steady state test the choice of  $\theta = 1$  is perfectly reasonable, since the problem is stationary. The results for the errors and for the convergence rates are reported in the following convergence table and it is possible to state that the semi-implicit DG method on staggered meshes is arbitrary high order accurate in space. In particular for a given polynomial degree  $P$ , we observe that the convergence rate is equal to  $P + 1$  for the velocity norm ( $V = \sqrt{u^2 + v^2}$ ). For the pressure, we notice an optimal convergence rate for  $P$  even and a sub-optimal convergence order for  $P$  odd. We attribute this behaviour to the choice of the basis functions; more detailed analyses will be carried out in the future.

Later on, we impose the components of the velocity  $[U_\infty, V_\infty]$  equal to  $[1, 1]$ . The computational grid is composed of 625 square cells with  $N_x = N_y = 25$ , the degree of the polynomial basis functions is  $P = 5$  and  $\Omega_{xy} = [-5, +5] \times [-5, +5]$ . We impose the implicitness parameter equal to  $\theta = 0.7$  and periodic boundary conditions are prescribed at the all the boundaries. Hence, at the final time  $t_{end} = 10$  the vortex is centred again in the initial position at  $[x_0, y_0] = [0, 0]$  and the exact solution corresponds to the initial condition reported in equations (5.68) and (5.69). In Fig. 5.19 we report the contour plots for the density and for the pressure where we observe that the

$P$	$N_x = N_y$	$\epsilon_{L_2}^P$	$\epsilon_{L_2}^{V_t}$	$O_{L_2}^P$	$O_{L_2}^{V_t}$
0	100	1.94E-2	3.90E-2	-	-
0	200	9.69E-3	1.71E-2	0.99	1.18
0	300	6.58E-3	1.10E-2	0.95	1.09
0	400	4.99E-3	8.10E-3	0.95	1.06
1	100	2.44E-2	3.14E-3	-	-
1	200	1.25E-2	9.83E-4	0.96	1.67
1	300	8.40E-3	4.64E-4	0.98	1.85
1	350	7.21E-3	3.52E-4	0.99	1.77
1	400	6.32E-3	2.77E-4	0.99	1.80
2	50	3.43E-3	1.67E-3	-	-
2	100	6.31E-4	2.76E-4	2.44	2.59
2	150	1.79E-4	9.03E-5	3.10	2.75
2	200	7.69E-5	3.88E-5	2.93	2.93
3	25	8.89E-4	2.90E-4	-	-
3	50	1.14E-4	2.49E-5	2.96	3.53
3	75	3.35E-5	6.36E-6	3.02	3.36
3	100	1.41E-5	2.51E-6	3.01	3.23
4	15	9.82E-4	3.42E-4	-	-
4	30	7.67E-5	2.14E-5	3.67	4.00
4	45	1.73E-5	3.04E-6	3.66	4.86
4	60	5.58E-6	1.03E-6	3.93	3.75
5	15	8.39E-4	3.35E-4	-	-
5	30	4.97E-5	5.27E-6	4.07	5.98
5	45	7.02E-6	5.89E-7	4.82	5.40
5	60	1.46E-6	2.85E-7	5.44	2.52

**Table 5.2:** Shu-Osher vortex - Numerical convergence rates computed with respect to the  $L_2$  error norms of pressure and velocity for the two-dimensional semi-implicit staggered DG scheme.

numerical solution is symmetric and clean because wiggles are not generated.



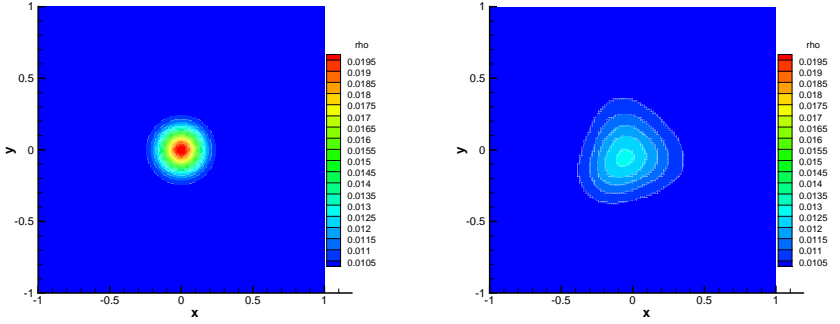
**Figure 5.19:** Isentropic vortex - Contour plots for the density and the pressure at time  $t = 10$ .

### 5.3.2 Advection of a smooth density bell in 2D

Similarly to the 1D method in section 5.2.1, this test consists in a smooth two dimensional Gaussian bell moving in a uniform pressure and velocity flow. Hence the initial condition reads as follows

$$\rho(x, y, 0) = \rho_0(1 + e^{-\frac{0.5r^2}{0.1^2}}) \quad p(x, y, 0) = 1 \quad u(x, y, 0) = 2 \quad v(x, y, 0) = 2. \quad (5.70)$$

The computational domain is  $\Omega_{xy} = [-1, +1] \times [-1, +1]$  and  $t_{end} = 1$ ; consequently imposing periodic boundary conditions the reference solution at the final time coincides with the initial condition in eq. (5.70). We consider two configurations both with  $\theta = 0.55$ ; in the first one the polynomial degree is  $P = 5$  and  $N_x = N_y = 50$  and for the second case  $N_x = N_y = 300$  using  $P = 0$ , which corresponds to the piecewise constant finite volume data. Note that in both situations the number of degrees of freedom is the same and equal to  $N_x \times N_y \times (P + 1)^2 = 9 \times 10^4$ . The results for the contour plot at time  $t = 1$  are shown in Figure 5.20. Here we observe that for the high order DG case the solution is centred and symmetric, while we can see that the effects of the numerical viscosity are relevant in the first order situation. Moreover, in Figure 5.21, we report a comparison between the analytical and the exact solutions both sliced along the  $x$  axis. Similarly, to the 1D case in subsection 5.2.1 in the FV case the dissipation is very high while the semi-implicit staggered DG scheme produces a very accurate solution. Consequently, we underline the importance of the high order approach for the numerical solution of nonlinear hyperbolic partial differential equations.



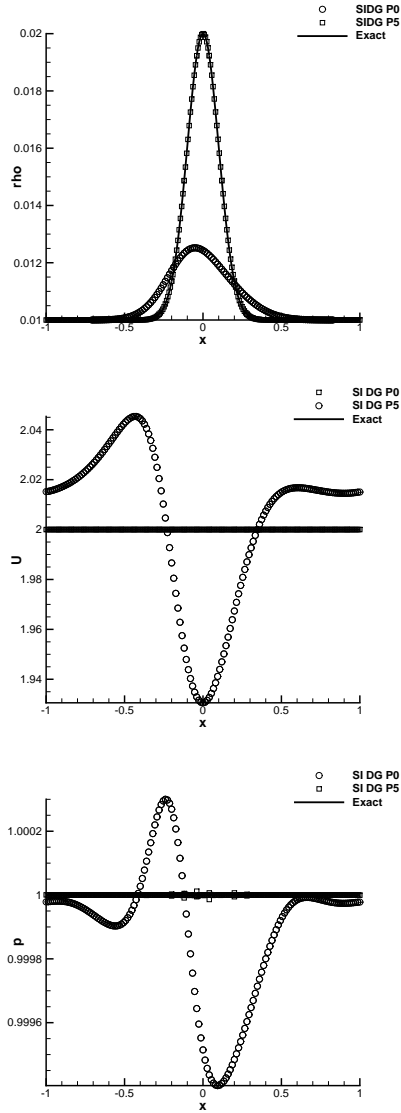
**Figure 5.20:** Advection of a smooth density bell in 2D - Density for  $P = 5$  (high order DG) and for  $P = 0$  (first order FV) at time  $t = 1$ .

### 5.3.3 Smooth two-dimensional acoustic wave propagation

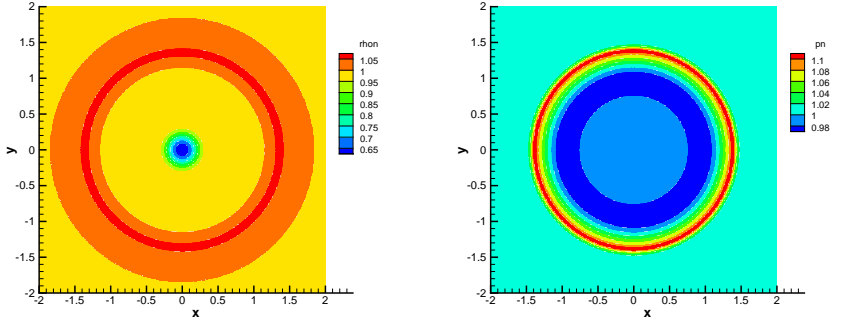
We consider another smooth test for the two-dimensional Euler equations. For the pressure the initial condition is a smooth surface,  $p(x, y, 0) = 1 + e^{-\alpha r^2}$ , while the density is constant,  $\rho(x, y, 0) = 1$ , and both components of velocity are null,  $u(x, y, 0) = 0$ ,  $v(x, y, 0) = 0$ . This simulation is characterized by very low Mach number however the stability condition of the semi-implicit DG scheme is based on the fluid velocity and consequently this scheme is very efficient for these kind of acoustic regimes. On the contrary, the same test solved numerically by an explicit method would be very inefficient; in fact, for this last case the stability condition is based on the maximum eigenvalues  $|u|+c$  and  $|v|+c$ .

For this simulation the computational domain is  $\Omega_{xy} = [-2, +2] \times [-2, +2]$  and we use  $N_x = N_y = 25$  with a total number of 625 cells using a polynomial degree  $P$  equal to 5. Hence, the total number of degrees of freedom is 22500. The implicitness parameter  $\theta$  is chosen equal to 0.55 since we want to reduce the error in time as much as possible.

In Fig. 5.22 we can observe that the contour graphics for the density and the pressure are perfectly clean and symmetric. In addition, in Fig. 5.23 we compare the numerical results against a reference solution obtained running a second order explicit TVD scheme on a very fine mesh. Except for some numerical viscosity close the to peaks, the agreement is excellent and all the acoustic waves travel with the proper speed. According to [140], we can state that the semi-implicit staggered DG scheme is a very suitable numerical method for CFD simulation in the low Mach number regime.



**Figure 5.21:** Advection of a smooth density bell in 2D - Reference solutions and numerical solutions along the  $x$  axis for  $\rho$ ,  $U$  and  $p$  at time  $t = 1$ .



**Figure 5.22:** Two-dimensional smooth acoustic wave propagation - Contour plots for the density and the pressure at time  $t = 1$ .

### 5.3.4 Circular explosion

Here we consider a well known two dimensional Riemann problem characterized by circular symmetry. Given a computational domain  $\Omega_{xy} = [-1, +1] \times [-1, +1]$  the initial condition reads

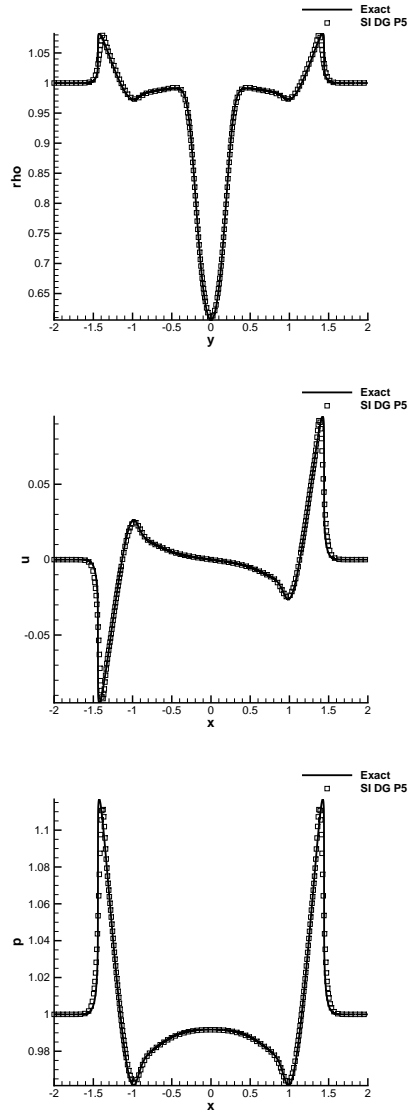
$$\mathbf{V}(x, y, 0) = \begin{cases} \mathbf{V}_{in} & \text{if } r \leq r_c \\ \mathbf{V}_{out} & \text{if } r > r_c. \end{cases} \quad (5.71)$$

where  $r = \sqrt{x^2 + y^2}$  is the radial coordinate and  $\mathbf{V} = [\rho, u, v, p]$  is the vector of state in primitive variables. For this simulation we impose the polynomial degree  $P$  equal to 5 and we use  $N_x = N_y = 100$  with  $\theta_{DG} = 0.55$  and  $t_{end} = 0.2$ . In Figure 5.24 we see that the density has a very clean contour plot. In addition we report a map of the troubled cells; most of them are located in the vicinity of the shock wave and other few control volumes are limited in the zone of the contact wave. Moreover, we take a cut of the solution along the  $x$  axis and we report a comparison against the reference solution in Figure 5.25. This last data are obtained using a robust second order TVD scheme (see [157]) on a very fine grid in order to solve an equivalent 1D Euler system of PDEs with a source term that considers properly the effect of the radial symmetry [148]. It is possible to observe a very good agreement between the data; furthermore in the numerical solution oscillations are not present and all the waves are computed with the correct speed of propagation.

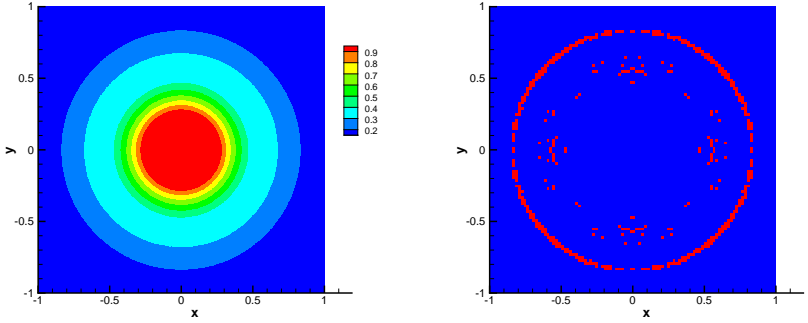
### 5.3.5 Two dimensional Riemann problems

In the last benchmark we consider a class of two dimensional Riemann problems presented in [129] and further studied in [102]. The computational domain  $\Omega_{xy} = [-0.5, +0.5] \times [-0.5, +0.5]$  is





**Figure 5.23:** Two-dimensional smooth acoustic wave propagation - Reference solutions and numerical solutions along the  $x$  axis for  $\rho$ ,  $U$  and  $p$  at time  $t = 1$ . 179



**Figure 5.24:** Circular explosion - Contour plots for the density and map of the troubled zones (red cell) at time  $t = 0.2$ .

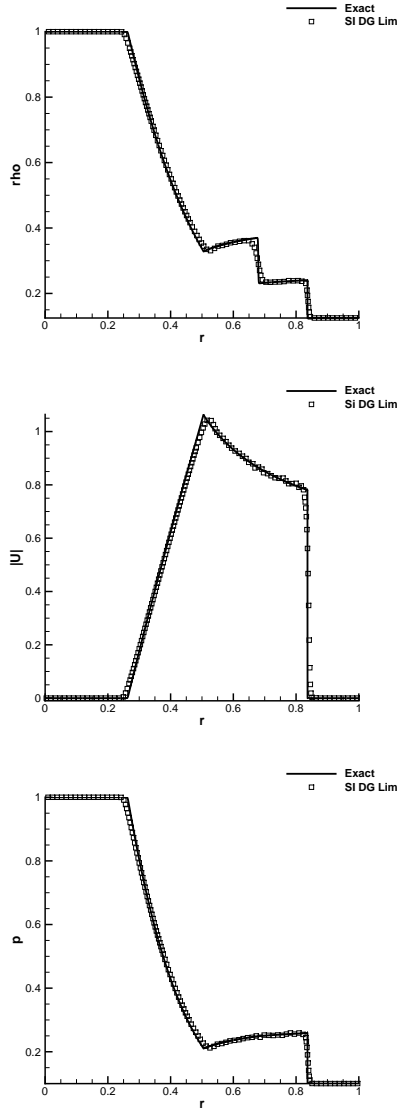
discretized using  $N_x = N_y = 50$  cells with a polynomial degree  $P = 5$ . The initial condition reads as follows

$$(\rho, u, v, p) = \begin{cases} (\rho_1, u_1, v_1, p_1) & \text{if } x > 0 \wedge y > 0, \\ (\rho_2, u_2, v_2, p_2) & \text{if } x \leq 0 \wedge y > 0, \\ (\rho_3, u_3, v_3, p_3) & \text{if } x \leq 0 \wedge y \leq 0, \\ (\rho_4, u_4, v_4, p_4) & \text{if } x > 0 \wedge y \leq 0. \end{cases} \quad (5.72)$$

We consider four configurations, the same of the simulations carried out in [55], and the parameters for the initial condition are listed in Table 5.4. From Figures 5.26-5.29 we can state that the numerical results are in good agreement with the simulations carried out in [55, 102]. In addition, we observe that most of the troubled control volumes are located on the discontinuities but, probability due to numerical noise, a few bad cells are also detected in the plateau regions and at the boundaries. Moreover, since this method is at most second order accurate in time, we observe a relevant numerical viscosity which is clearly higher here than in WENO-based finite volume limiters of explicit DG schemes, see [64, 163].

## 5.4 Conclusions about the subcell finite volume limiter for SIDG schemes applied to the compressible Euler equations

We proposed a novel family of semi-implicit DG schemes on staggered meshes with a *a posteriori* sub-cell finite volume limiter applied to the 1D and 2D Euler equations of compressible gasdynamics. In particular, we extended the FV method presented in [55] to high order of accuracy.

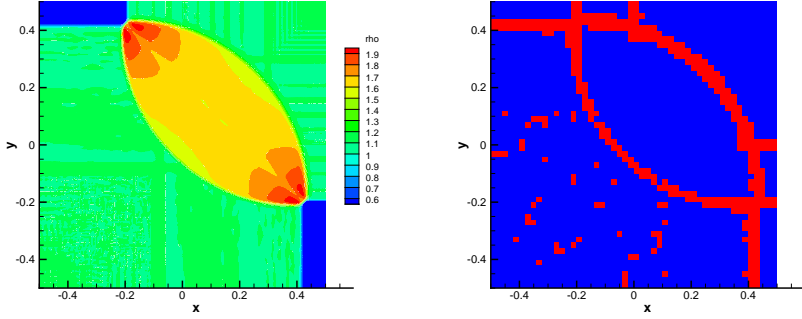


**Figure 5.25:** Circular explosion - Reference solutions and cut of the numerical solutions along the  $x$  axis for  $\rho, U$  and  $p$  at time  $t = 0.2$ .

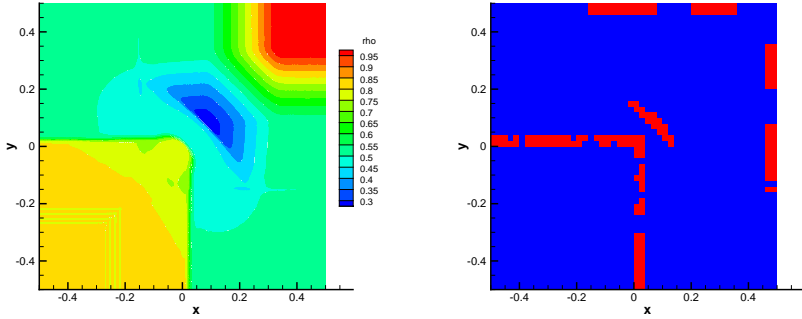
**Table 5.3:** Initial conditions for the two-dimensional Riemann problems.

Configuration C4				
	$\rho$	$U$	$V$	$p$
$(x > 0, y > 0)$	1.1	0.0	0.0	1.1
$(x < 0, y > 0)$	0.5065	0.8939	0.0	0.35
$(x < 0, y < 0)$	1.1	0.8939	0.8939	1.1
$(x > 0, y < 0)$	0.5065	0.0	0.8939	0.35
Configuration C7				
	$\rho$	$U$	$V$	$p$
$(x > 0, y > 0)$	0.5197	-0.6259	0.1	0.4
$(x < 0, y > 0)$	1.0	0.1	0.1	1.0
$(x < 0, y < 0)$	0.8	0.1	0.1	0.4
$(x > 0, y < 0)$	0.5197	0.1	-0.6259	0.4
Configuration C8				
	$\rho$	$U$	$V$	$p$
$(x > 0, y > 0)$	0.5197	0.1	0.1	0.4
$(x < 0, y > 0)$	1.0	-0.6259	0.1	1.0
$(x < 0, y < 0)$	0.8	0.1	0.1	1.0
$(x > 0, y < 0)$	1.0	0.1	-0.6259	1.0
Configuration C16				
	$\rho$	$U$	$V$	$p$
$(x > 0, y > 0)$	0.5313	0.1	0.1	0.4
$(x < 0, y > 0)$	1.0222	-0.6179	0.1	1.0
$(x < 0, y < 0)$	0.8	0.1	0.1	1.0
$(x > 0, y < 0)$	1.0	0.1	0.8276	1.0

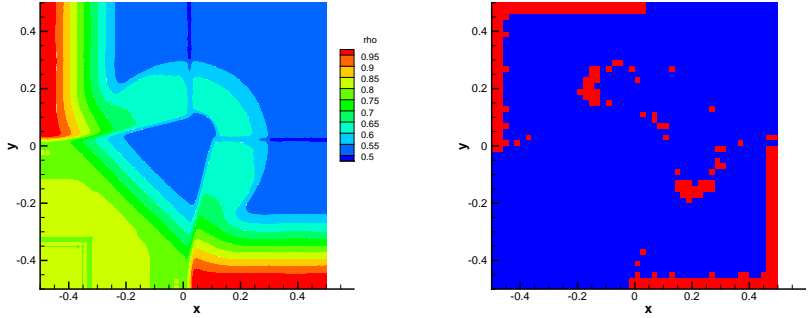
**Table 5.4:** Initial condition for the 2D Riemann problem



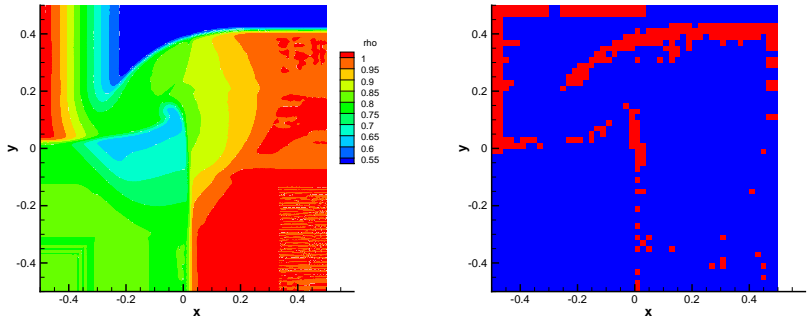
**Figure 5.26:** Two dimensional Riemann problem RP2D C4 - Contour plots for the density and map of the troubled zones (red cell) at time  $t = 0.2$ .



**Figure 5.27:** Two dimensional Riemann problem RP2D C7 - Contour plots for the density and map of the troubled zones (red cell) at time  $t = 0.2$ .



**Figure 5.28:** Two dimensional Riemann problem RP2D C8 - Contour plots for the density and map of the troubled zones (red cell) at time  $t = 0.2$ .



**Figure 5.29:** Two dimensional Riemann problem RP2D C16 - Contour plots for the density and map of the troubled zones (red cell) at time  $t = 0.2$ .

Subsequently, in order to develop a robust shock-capturing strategy, we followed the approach used in Chapter 4, where the *a posteriori* subcell finite volume limiter was extended to staggered semi-implicit schemes for the first time.

Similarly to the other *a posteriori* sub-cell finite volume limiters for explicit DG schemes on collocated grids developed in [14, 59, 64], this approach is based on the MOOD paradigm introduced in [39, 47, 48, 107]. Hence, an unlimited semi-implicit staggered DG method is first applied in order to produce a so-called *candidate solution* at time  $t^{n+1}$ . Then, applying physical and numerical admissibility criteria *troubled cells* characterized by a non-valid solution are detected. Then, a more robust first order semi-implicit staggered finite volume scheme is applied in the control volumes flagged as troubled cells. Successively, the linear system for the pressure is solved again involving the unlimited DG cells and the finite volume sub-cells. The algorithm is concluded with the reconstruction of the DG polynomial from the piecewise constant subcell averages within the troubled zones.

Several benchmarks have been carried out in order to confirm that the new schemes behave well both for the low Mach number regime, due to the implicit treatment of the pressure term, and both for several kinds of Riemann problem, since that the *a posteriori* subcell finite volume limiter is able to stabilize the DG scheme in the presence of shock waves and other flow discontinuities.





## 6 Conclusions and Outlook

In this thesis we have investigated several new *semi-implicit* approaches for the numerical solution of non-linear systems of partial differential equations governing different phenomena of computational fluid dynamics (CFD). This class of methods is based on the idea of solving explicitly the non-linear convective terms of the system and of discretizing implicitly the remaining pressure-based part of the system.

In the first part of this PhD thesis, we have considered PDE models for weakly compressible viscous fluids in elastic pipes. In Chapter 2 we carried out a detailed comparison of fully explicit and semi-implicit 1D and 2D finite volume schemes for the simulation of highly unsteady pressurized laminar flows in flexible pipes. In particular we considered a cross-sectionally averaged 1D and a  $2D_{xr}$  model, which can be derived from the three dimensional Navier-Stokes equations under the assumption of axial symmetry and hydrostatic pressure. Furthermore, great attention was dedicated to the influence of the *frequency dependent friction*. We considered the two main families of unsteady friction models: *convolution integral* (CI) and *instantaneous acceleration* (IA) models. Our results have clearly shown that CI models are significantly better than IA models in terms of accuracy. In addition, for CI models the computation of the time and memory consuming convolution integral can be circumvented if the unsteady friction term is properly rewritten as a set of additional *Ordinary differential equations* for some auxiliary quantities. Moreover, from our simulations it is evident that, concerning the computational efficiency and the accuracy, semi-implicit finite volume schemes are significantly superior to explicit finite volume methods in the context of compressible pipe flow. In Chapter 3 we implemented high order accurate semi-implicit Discontinuous Galerkin schemes on *staggered grids* for the 1D and the  $2D_{xr}$  pipe flow models. The schemes developed are characterized by arbitrary high order of accuracy in space and the  $\theta$ -method allows to go up to second order of accuracy in time. Moreover, due to the implicit treatment of the pressure, for these schemes the time step  $\Delta t$  is given by a *mild stability condition* that is based only on the fluid velocity and not on the sound speed. In both schemes, the pressure is obtained after the solution of *mildly non-linear systems* using the *Newton-type algorithm* of Brugnano and Casulli [18, 19]. Moreover, since we adopted a staggered grid strategy, we observed that the system is symmetric, block three-diagonal and well conditioned. Several benchmarks have been carried out in order to validate the novel schemes. In all cases we observed an excellent agreement between numerical simulation and reference solution.

Later on we developed a novel *a posteriori* subcell finite volume limiter for high-order staggered semi-implicit DG schemes. In particular, in our opinion this tool is fundamental in order to develop further extensions of staggered semi-implicit DG to real world problems. In order to avoid *Gibbs phenomena* in high order discontinuous finite elements methods, the *a posteriori* FV subcell limiting is a successful strategy, proposed in [59, 64, 163], and based on the MOOD algorithm of Clain, Loubère and Diot [39, 47, 48, 107]. In this PhD thesis, it was applied for the first time to semi-implicit staggered DG methods. Considering the *shallow water equations* (see Chapter 4), we first use the unlimited semi-implicit DG scheme in order to produce a so-called *candidate solution*

at time  $t^{n+1}$ . Successively, the validity of this solution is analysed using *physical and numerical admissibility criteria*. The cells that do not have an admissible solution are marked as *troubled*. Then, the linear system is assembled again, considering the unlimited DG in the cells that are not troubled and using a more robust first order finite volume subcell method in the troubled control volumes. This kind of limiting is further extended to the Euler equations of compressible gasdynamics in Chapter 5. The method is validated for several stringent Riemann problems and other benchmarks that involve strong discontinuities and shock waves. The scheme works very well in all these test cases and spurious oscillations are not produced. In addition, the troubled cells are typically located in the zones close to the discontinuities and, in general, they are not detected for problems characterized by smooth solutions.

In future work, we will consider the extension of the semi-implicit finite volume and DG method to more complex nonlinear systems of PDEs, for example the Baer-Nunziato [1] model for compressible multiphase flows as well as the unified first order hyperbolic Godunov-Peshkov-Romenski (GPR) model of continuum mechanics proposed in [60, 119].

Further developments may also regard the extension and implementation of the novel *a posteriori* subcell FV limiter for high order DG schemes for the compressible Navier-Stokes equations, see [55, 140], and for the MHD equations, see [66]. In addition, we think that also the Richards equation, see [30, 31], could be solved by the semi-implicit staggered DG scheme, and in this framework the *a posteriori* limiter could be used for solutions with very sharp fronts or steep gradients, or in order to deal with discontinuities in the hydraulic conductivity within the soil.

Finally, in the framework of pipe flow simulations, future advances will concern the implementation of more complicated equations of state (EOS) in order to model phase transitions as well as the study of the effects of the full compressibility of the fluid, including the energy equation. Other improvements will regard the development of elastic models that are governed by PDEs, see [104], in order to describe properly the phenomena of fluid-structure interaction.

# A Appendices

## A.1 Matrices and tensors for staggered semi-implicit schemes for the 1D shallow water equations

### A.1.1 Matrices and tensors for the 1D semi-implicit DG method

$$\mathbf{M} = \int_0^1 \varphi(\xi) \varphi(\xi) d\xi \quad (\text{A.1})$$

$$\mathbf{K} = \int_0^1 \varphi'(\xi) \varphi(\xi) d\xi \quad (\text{A.2})$$

$$\mathbf{R}_J^{\text{DG}} = \varphi(1) \varphi\left(\frac{1}{2}\right) - \int_{\frac{1}{2}}^1 \varphi'(\xi) \varphi\left(\xi - \frac{1}{2}\right) d\xi \quad (\text{A.3})$$

$$\mathbf{L}_J^{\text{DG}} = \varphi(0) \varphi\left(\frac{1}{2}\right) + \int_0^{\frac{1}{2}} \varphi'(\xi) \varphi\left(\xi + \frac{1}{2}\right) d\xi \quad (\text{A.4})$$

$$\mathbf{R}_u^{\text{DG}} = \varphi\left(\frac{1}{2}\right) \varphi\left(\frac{1}{2}\right) \varphi(0) + \int_{\frac{1}{2}}^1 \varphi(\xi) \varphi(\xi) \varphi'\left(\xi - \frac{1}{2}\right) d\xi \quad (\text{A.5})$$

$$\mathbf{L}_u^{\text{DG}} = \varphi\left(\frac{1}{2}\right) \varphi\left(\frac{1}{2}\right) \varphi(1) - \int_0^{\frac{1}{2}} \varphi(\xi) \varphi(\xi) \varphi'\left(\xi + \frac{1}{2}\right) d\xi \quad (\text{A.6})$$

$$\mathbf{M}_L^{\text{DG}} = \int_0^{\frac{1}{2}} \varphi(\xi) \varphi\left(\xi + \frac{1}{2}\right) d\xi \quad (\text{A.7})$$

$$\mathbf{M}_R^{\text{DG}} = \int_{\frac{1}{2}}^1 \varphi(\xi) \varphi\left(\xi - \frac{1}{2}\right) d\xi \quad (\text{A.8})$$

### A.1.2 Matrices and tensors for the 1D semi-implicit sub-cell FV method for P=2

$$\mathbf{R}^{\text{FV}} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 5 & 0 & 0 & 0 & 0 \\ -5 & 5 & 0 & 0 & 0 \\ 0 & -5 & 5 & 0 & 0 \end{pmatrix} \quad \mathbf{L}^{\text{FV}} = \begin{pmatrix} 0 & 0 & 5 & -5 & 0 \\ 0 & 0 & 0 & 5 & -5 \\ 0 & 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad (\text{A.9})$$

$$\mathbf{M}_{\mathbf{R}}^{\text{FV}} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0.5 & 0 & 0 & 0 & 0 \\ 0.5 & 0.5 & 0 & 0 & 0 \\ 0 & 0.5 & 0.5 & 0 & 0 \end{pmatrix} \quad \mathbf{M}_{\mathbf{L}}^{\text{FV}} = \begin{pmatrix} 0 & 0 & 0.5 & 0.5 & 0 \\ 0 & 0 & 0 & 0.5 & 0.5 \\ 0 & 0 & 0 & 0 & 0.5 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad (\text{A.10})$$

## A.2 Matrices and tensors for staggered semi-implicit schemes for the 2D shallow water equations

### A.2.1 Matrices and tensors for the 2D semi-implicit DG method

$$\mathbf{M} = \int_0^1 \int_0^1 \varphi(\xi, \gamma) \varphi(\xi, \gamma) d\xi d\gamma \quad (\text{A.11})$$

$$\mathbf{R}_{\mathbf{j}}^{\text{x,DG}} = \int_0^1 \varphi(1, \gamma) \varphi\left(\frac{1}{2}, \gamma\right) d\gamma - \int_{\frac{1}{2}}^1 \int_0^1 \frac{\partial \varphi(\xi, \gamma)}{\partial \xi} \varphi\left(\xi - \frac{1}{2}, \gamma\right) d\gamma d\xi \quad (\text{A.12})$$

$$\mathbf{R}_{\mathbf{j}}^{\text{y,DG}} = \int_0^1 \varphi(\xi, 1) \varphi\left(\xi, \frac{1}{2}\right) d\xi - \int_0^{\frac{1}{2}} \int_{\frac{1}{2}}^1 \frac{\partial \varphi(\xi, \gamma)}{\partial \gamma} \varphi\left(\xi, \gamma - \frac{1}{2}\right) d\gamma d\xi \quad (\text{A.13})$$

$$\mathbf{L}_{\mathbf{j}}^{\text{x,DG}} = \int_0^1 \varphi(0, \gamma) \varphi\left(\frac{1}{2}, \gamma\right) d\gamma + \int_0^{\frac{1}{2}} \int_0^1 \frac{\partial \varphi(\xi, \gamma)}{\partial \xi} \varphi\left(\xi + \frac{1}{2}, \gamma\right) d\gamma d\xi \quad (\text{A.14})$$

$$\mathbf{L}_{\mathbf{j}}^{\text{y,DG}} = \int_0^1 \varphi(\xi, 0) \varphi\left(\xi, \frac{1}{2}\right) d\xi + \int_0^{\frac{1}{2}} \int_0^1 \frac{\partial \varphi(\xi, \gamma)}{\partial \gamma} \varphi\left(\xi, \gamma + \frac{1}{2}\right) d\gamma d\xi \quad (\text{A.15})$$

$$\mathbf{R}_{\mathbf{u}}^{\text{x,DG}} = \int_0^1 \varphi\left(\frac{1}{2}, \gamma\right) \varphi\left(\frac{1}{2}, \gamma\right) \varphi(0, \gamma) d\gamma + \int_{\frac{1}{2}}^1 \int_0^1 \varphi(\xi, \gamma) \varphi(\xi, \gamma) \frac{\partial \varphi\left(\xi - \frac{1}{2}, \gamma\right)}{\partial \xi} d\gamma d\xi \quad (\text{A.16})$$

$$\mathbf{R}_u^{y,DG} = \int_0^1 \varphi\left(\xi, \frac{1}{2}\right) \varphi\left(\xi, \frac{1}{2}\right) \varphi(\xi, 0) d\xi + \int_0^1 \int_{\frac{1}{2}}^1 \varphi(\xi, \gamma) \varphi(\xi, \gamma) \frac{\partial \varphi\left(\xi, \gamma - \frac{1}{2}\right)}{\partial \gamma} d\gamma d\xi \quad (\text{A.17})$$

$$\mathbf{L}_u^{x,DG} = \int_0^1 \varphi\left(\frac{1}{2}, \gamma\right) \varphi\left(\frac{1}{2}, \gamma\right) \varphi(1, \gamma) d\gamma - \int_0^{\frac{1}{2}} \int_0^1 \varphi(\xi, \gamma) \varphi(\xi, \gamma) \frac{\partial \varphi\left(\xi + \frac{1}{2}, \gamma\right)}{\partial \xi} d\gamma d\xi \quad (\text{A.18})$$

$$\mathbf{L}_u^{y,DG} = \int_0^1 \varphi\left(\xi, \frac{1}{2}\right) \varphi\left(\xi, \frac{1}{2}\right) \varphi(\xi, 1) d\xi - \int_0^1 \int_0^{\frac{1}{2}} \varphi(\xi, \gamma) \varphi(\xi, \gamma) \frac{\partial \varphi\left(\xi, \gamma + \frac{1}{2}\right)}{\partial \gamma} d\gamma d\xi \quad (\text{A.19})$$

$$\mathbf{M}_L^{x,DG} = \int_0^{\frac{1}{2}} \int_0^1 \varphi(\xi, \gamma) \varphi\left(\xi + \frac{1}{2}, \gamma\right) d\xi d\gamma \quad (\text{A.20})$$

$$\mathbf{M}_L^{y,DG} = \int_0^1 \int_0^{\frac{1}{2}} \varphi(\xi, \gamma) \varphi\left(\xi, \gamma + \frac{1}{2}\right) d\xi d\gamma \quad (\text{A.21})$$

$$\mathbf{M}_R^{x,DG} = \int_{\frac{1}{2}}^1 \int_0^1 \varphi(\xi, \gamma) \varphi\left(\xi - \frac{1}{2}, \gamma\right) d\xi d\gamma \quad (\text{A.22})$$

$$\mathbf{M}_R^{y,DG} = \int_0^1 \int_{\frac{1}{2}}^1 \varphi(\xi, \gamma) \varphi\left(\xi, \gamma - \frac{1}{2}\right) d\xi d\gamma \quad (\text{A.23})$$

### A.2.2 Matrices and tensors for the 2D semi-implicit sub-cell FV method for P=1

$$\mathbf{R}^{\mathbf{x},\text{FV}} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -3 & 3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -3 & 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -3 & 3 & 0 \end{pmatrix} \quad (\text{A.24})$$

$$\mathbf{L}^{\mathbf{x},\text{FV}} = \begin{pmatrix} 0 & 3 & -3 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 & -3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & -3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad (\text{A.25})$$

$$\mathbf{R}^{\mathbf{y},\text{FV}} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 & 0 & 0 & 0 & 0 \\ -3 & 0 & 0 & 3 & 0 & 0 & 0 & 0 & 0 \\ 0 & -3 & 0 & 0 & 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & -3 & 0 & 0 & 3 & 0 & 0 & 0 \end{pmatrix} \quad (\text{A.26})$$

$$\mathbf{L}^{\mathbf{y},\text{FV}} = \begin{pmatrix} 0 & 0 & 0 & 3 & 0 & 0 & -3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 & 0 & 0 & -3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 3 & 0 & 0 & -3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad (\text{A.27})$$

$$\mathbf{M}_R^{\mathbf{x},\text{FV}} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.5 & 0.5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.5 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.5 & 0.5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.5 & 0.5 & 0 \end{pmatrix} \quad (\text{A.28})$$

$$\mathbf{M}_L^{\mathbf{x},\text{FV}} = \begin{pmatrix} 0 & 0.5 & 0.5 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.5 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.5 & 0.5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.5 & 0.5 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.5 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad (\text{A.29})$$

$$\mathbf{M}_R^{\mathbf{y},\text{FV}} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.5 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.5 & 0 & 0 & 0.5 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.5 & 0 & 0 & 0.5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.5 & 0 & 0 & 0.5 & 0 & 0 & 0 \end{pmatrix} \quad (\text{A.30})$$

$$\mathbf{M}_L^{\mathbf{y},\text{FV}} = \begin{pmatrix} 0 & 0 & 0 & 0.5 & 0 & 0 & 0.5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.5 & 0 & 0 & 0.5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.5 & 0 & 0 & 0.5 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.5 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad (\text{A.31})$$

### A.3 Matrices and tensors for the 1D semi-implicit schemes for Euler equation

#### A.3.1 Matrices and tensors for the 1D semi-implicit DG method

$$\mathbf{M} = \int_0^1 \varphi(\xi) \varphi(\xi) d\xi \quad (\text{A.32})$$

$$\mathbf{K} = \int_0^1 \varphi'(\xi) \varphi(\xi) d\xi \quad (\text{A.33})$$

$$\mathbf{R}_e^{\text{DG}} = \varphi(1) \varphi\left(\frac{1}{2}\right) \varphi\left(\frac{1}{2}\right) - \int_{\frac{1}{2}}^1 \varphi'(\xi) \varphi\left(\xi - \frac{1}{2}\right) \varphi\left(\xi - \frac{1}{2}\right) d\xi \quad (\text{A.34})$$

$$\mathbf{L}_e^{\text{DG}} = \varphi(0) \varphi\left(\frac{1}{2}\right) \varphi\left(\frac{1}{2}\right) + \int_0^{\frac{1}{2}} \varphi'(\xi) \varphi\left(\xi + \frac{1}{2}\right) \varphi\left(\xi + \frac{1}{2}\right) d\xi \quad (\text{A.35})$$

$$\mathbf{R}_p^{\text{DG}} = \varphi\left(\frac{1}{2}\right) \varphi(0) + \int_{\frac{1}{2}}^1 \varphi(\xi) \varphi'\left(\xi - \frac{1}{2}\right) d\xi \quad (\text{A.36})$$

$$\mathbf{L}_p^{\text{DG}} = \varphi\left(\frac{1}{2}\right) \varphi(1) - \int_0^{\frac{1}{2}} \varphi(\xi) \varphi'\left(\xi + \frac{1}{2}\right) d\xi \quad (\text{A.37})$$

$$\mathbf{M}_L^{\text{DG}} = \int_0^{\frac{1}{2}} \varphi(\xi) \varphi\left(\xi + \frac{1}{2}\right) d\xi \quad (\text{A.38})$$

$$\mathbf{M}_R^{\text{DG}} = \int_{\frac{1}{2}}^1 \varphi(\xi) \varphi\left(\xi - \frac{1}{2}\right) d\xi \quad (\text{A.39})$$

#### A.3.2 Matrices and tensors for the 1D semi-implicit sub-cell FV method for P=2 for Euler equation

$$\mathbf{R}^{\text{FV}} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 5 & 0 & 0 & 0 & 0 \\ -5 & 5 & 0 & 0 & 0 \\ 0 & -5 & 5 & 0 & 0 \end{pmatrix} \quad \mathbf{L}^{\text{FV}} = \begin{pmatrix} 0 & 0 & 5 & -5 & 0 \\ 0 & 0 & 0 & 5 & -5 \\ 0 & 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad (\text{A.40})$$



$$\mathbf{M}_R^{\text{FV}} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0.5 & 0 & 0 & 0 & 0 \\ 0.5 & 0.5 & 0 & 0 & 0 \\ 0 & 0.5 & 0.5 & 0 & 0 \end{pmatrix} \quad \mathbf{M}_L^{\text{FV}} = \begin{pmatrix} 0 & 0 & 0.5 & 0.5 & 0 \\ 0 & 0 & 0 & 0.5 & 0.5 \\ 0 & 0 & 0 & 0 & 0.5 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad (\text{A.41})$$

### A.3.3 Tensors for the limited 1D semi-implicit DG method for Euler equation

$$\mathbf{R}_e^{\text{DG}} = \varphi(1)\varphi\left(\frac{1}{2}\right) - \int_{\frac{1}{2}}^1 \varphi'(\xi)\varphi\left(\xi - \frac{1}{2}\right) d\xi \quad (\text{A.42})$$

$$\mathbf{L}_e^{\text{DG}} = \varphi(0)\varphi\left(\frac{1}{2}\right) + \int_0^{\frac{1}{2}} \varphi'(\xi)\varphi\left(\xi + \frac{1}{2}\right) d\xi \quad (\text{A.43})$$



# Bibliography

- [1] M.R. Baer and J.W. Nunziato. A two-phase mixture theory for the deflagration-to-detonation transition (ddt) in reactive granular materials. *International Journal of Multiphase Flow*, 12(6):861 – 889, 1986.
- [2] D. Balsara, C. Altmann, C.D. Munz, and M. Dumbser. A sub-cell based indicator for troubled zones in RKDG schemes and a novel class of hybrid RKDG+HWENO schemes. *Journal of Computational Physics*, 226:586–620, 2007.
- [3] G.E. Barter and D.L. Darmofal. Shock Capturing with PDE-based Artificial Viscosity for DGFEM: Part I. Formulation. *Journal of Computational Physics*, 229(5):1810–1827, 2010.
- [4] F. Bassi, L. Botti, A. Colombo, A. Ghidoni, and F.C. Massa. Linearly implicit Rosenbrock-type Runge-Kutta schemes applied to the Discontinuous Galerkin solution of compressible and incompressible unsteady flows. *Computers and Fluids*, 118:305–320, 2015.
- [5] F. Bassi, A. Crivellini, D.A. Di Pietro, and S. Rebay. An artificial compressibility flux for the discontinuous Galerkin solution of the incompressible Navier–Stokes equations. *Journal of Computational Physics*, 218:208–221, 2006.
- [6] F. Bassi, A. Crivellini, D.A. Di Pietro, and S. Rebay. An implicit high-order discontinuous Galerkin method for steady and unsteady incompressible flows. *Computers and Fluids*, 36:1529–1546, 2007.
- [7] F. Bassi and S. Rebay. A high-order accurate discontinuous finite element method for the numerical solution of the compressible Navier-Stokes equations. *Journal of Computational Physics*, 131:267–279, 1997.
- [8] C. Baumann and J. Oden. A discontinuous hp finite element method for convection-diffusion problems. *Computer Methods in Applied Mechanics and Engineering*, 175:311–341, 1999.
- [9] C. Baumann and J. Oden. A discontinuous hp finite element method for the Euler and Navier-Stokes equation. *International Journal for Numerical Methods in Fluids*, 31:79–95, 1999.
- [10] A. Bermudez and M.E. Vázquez-Cendón. Upwind methods for hyperbolic conservation laws with source terms. *Computers and Fluids*, 23(8):1049 – 1071, 1994.
- [11] R. Bernetti, V. A. Titarev, and E. F. Toro. Exact solution of the Riemann problem for the shallow water equations with discontinuous bottom geometry. *Journal of Computational Physics*, 227:3212–3243, March 2008.

- [12] H. Blasius. Grenzsichten in Flüssigkeiten mit kleiner Reibung. *Z. Math. Physik*, 56:1–37, 1908.
- [13] L. Bonaventura and A. Della Rocca. Unconditionally Strong Stability Preserving Extensions of the TR-BDF2 Method. *Journal of Scientific Computing*, 70(2):859–895, Feb 2017.
- [14] W. Boscheri and M. Dumbser. Arbitrary-Lagrangian-Eulerian Discontinuous Galerkin schemes with a posteriori subcell finite volume limiting on moving unstructured meshes. *Journal of Computational Physics*, 346:449 – 479, 2017.
- [15] Walter Boscheri. An efficient high order direct ALE ADER finite volume scheme with a posteriori limiting for hydrodynamics and magnetohydrodynamics. *International Journal for Numerical Methods in Fluids*, 84(2):76–106, 2017.
- [16] Walter Boscheri and Michael Dumbser. Arbitrary-lagrangian-eulerian one-step weno finite volume schemes on unstructured triangular meshes. *Communications in Computational Physics*, 14(5):1174–1206, 2013.
- [17] Walter Boscheri and Michael Dumbser. A direct arbitrary-lagrangian-eulerian ader-weno finite volume scheme on unstructured tetrahedral meshes for conservative and non-conservative hyperbolic systems in 3d. *Journal of Computational Physics*, 275:484 – 523, 2014.
- [18] L. Brugnano and V. Casulli. Iterative solution of piecewise linear systems. *SIAM Journal on Scientific Computing*, 30:463–472, 2007.
- [19] L. Brugnano and V. Casulli. Iterative solution of piecewise linear systems and applications to flows in porous media. *SIAM Journal on Scientific Computing*, 31:1858–1873, 2009.
- [20] Luigi Brugnano and Vincenzo Casulli. Iterative solution of piecewise linear systems. *SIAM Journal on Scientific Computing*, 30(1):463–472, 2008.
- [21] Luigi Brugnano and Vincenzo Casulli. Iterative solution of piecewise linear systems and applications to flows in porous media. *SIAM Journal on Scientific Computing*, 31(3):1858–1873, 2009.
- [22] B. Brunone, U. M. Golia, and M. Greco. Modelling of fast transients by numerical methods. In *Proceeding of International Conference on Hydraulic Transients with water column separation*, pages 273–280, 1991.
- [23] B. Brunone, U. M. Golia, and M. Greco. Some remarks on the momentum equation for fast transients. In *Proceeding of International Conference on Hydraulic Transients with water column separation*, pages 201–209, 1991.
- [24] A. Burbeau, P. Sagaut, and Ch.-H. Bruneau. A Problem-Independent Limiter for High-Order Runge-Kutta Discontinuous Galerkin Methods. *Journal of Computational Physics*, 169(1):111 – 150, 2001.
- [25] C. Canuto, M.Y. Hussaini, and A. Quarteroni T.A. Zang. *Spectral Methods*. Springer, 2006.

- [26] Manuel Castro, José M. Gallardo, and Carlos Parés. High order finite volume schemes based on reconstruction of states for solving hyperbolic systems with nonconservative products. applications to shallow-water systems. *Mathematics of Computation*, 75(255):1103–1134, 2006.
- [27] V. Casulli. Semi-implicit finite difference methods for the two–dimensional shallow water equations. *Journal of Computational Physics*, 86:56–74, 1990.
- [28] V. Casulli. A high-resolution wetting and drying algorithm for free-surface hydrodynamics. *International Journal for Numerical Methods in Fluids*, 60:391–408, 2009.
- [29] V. Casulli. A semi–implicit numerical method for the free–surface Navier–Stokes equations. *International Journal for Numerical Methods in Fluids*, 74:605–622, 2014.
- [30] V. Casulli. A conservative semi-implicit method for coupled surface-subsurface flows in regional scale. *International Journal for Numerical Methods in Fluids*, 79(4):199–214, 2015. cited By 4.
- [31] V. Casulli. A coupled surface-subsurface model for hydrostatic flows under saturated and variably saturated conditions. *International Journal for Numerical Methods in Fluids*, 85(8):449–464, 2017. cited By 0.
- [32] V. Casulli and E. Cattani. Stability, accuracy and efficiency of a semi-implicit method for three-dimensional shallow water flow. *Computers and Mathematics with Applications*, 27(4):99 – 112, 1994.
- [33] V. Casulli and R. T. Cheng. Semi-implicit finite difference methods for three–dimensional shallow water flow. *International Journal for Numerical Methods in Fluids*, 15:629–648, 1992.
- [34] V. Casulli, M. Dumbser, and E. F. Toro. Semi-implicit numerical modeling of axially symmetric flows in compliant arterial systems. *International Journal for Numerical Methods in Biomedical Engineering*, 28:257–272, 2012.
- [35] V. Casulli and S. Stelling. Semi-implicit subgrid modelling of three-dimensional free-surface flows. *International Journal for Numerical Methods in Fluids*, 67:441–449, 2010.
- [36] V. Casulli and R. A. Walters. An unstructured grid, three–dimensional model based on the shallow water equations. *International Journal for Numerical Methods in Fluids*, 32:331–348, 2000.
- [37] V. Casulli and P. Zanolli. A nested Newton–type algorithm for finite volume methods solving Richards’ equation in mixed form. *SIAM Journal of Scientific Computing*, 32:2255–2273, 2009.
- [38] V. Casulli and P. Zanolli. Iterative solutions of mildly nonlinear systems. *Journal of Computational and Applied Mathematics*, 236:3937–3947, 2012.

- [39] S. Clain, S. Diot, and R. Loubère. A high-order finite volume method for systems of conservation laws-Multi-dimensional Optimal Order Detection (MOOD). *Journal of Computational Physics*, 230(10):4028 – 4050, 2011.
- [40] B. Cockburn, S.-Y. Lin, and C.-W. Shu. TVB Runge Kutta Local Projection Discontinuous Galerkin Finite Element Method for Conservation Laws III: One-Dimensional Systems. *Journal of Computational Physics*, 84:90, September 1989.
- [41] B. Cockburn, S. Y. Lin, and C. W. Shu. TVB Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws III: one dimensional systems. *Journal of Computational Physics*, 84:90–113, 1989.
- [42] B. Cockburn and C. W. Shu. TVB Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws II: general framework. *Mathematics of Computation*, 52:411–435, 1989.
- [43] B. Cockburn and C. W. Shu. The Runge-Kutta local projection P1-Discontinuous Galerkin finite element method for scalar conservation laws. *Mathematical Modelling and Numerical Analysis*, 25:337–361, 1991.
- [44] B. Cockburn and C. W. Shu. The Runge-Kutta discontinuous Galerkin method for conservation laws V: multidimensional systems. *Journal of Computational Physics*, 141:199–224, 1998.
- [45] P. Colella. A direct eulerian muscl scheme for gas dynamics. *SIAM Journal of Scientific and Statistical Computing*, 6:104 – 117, 1985.
- [46] R. Courant, K. Friedrichs, and H. Lewy.
- [47] S. Diot, S. Clain, and R. Loubère. Improved detection criteria for the Multi-dimensional Optimal Order Detection (MOOD) on unstructured meshes with very high-order polynomials. *Computers and Fluids*, 64(Supplement C):43 – 63, 2012.
- [48] S. Diot, R. Loubère, and S. Clain. The Multidimensional Optimal Order Detection method in the three-dimensional case: very high-order finite volume method for hyperbolic systems. *International Journal for Numerical Methods in Fluids*, 73(4):362–392, 2013.
- [49] D.Kuzmin. Slope limiting for discontinuous Galerkin approximations with a possibly non-orthogonal Taylor basis. *International Journal for Numerical Methods in Fluids*, 71(9):1178–1190, 2013.
- [50] D.Kuzmin. Hierarchical slope limiting in explicit and implicit discontinuous Galerkin method. *Journal of Computational Physics*, 257, Part B(0):1140 – 1162, 2014. Physics-compatible numerical methods.
- [51] V. Dolejsi and M. Feistauer. A semi-implicit discontinuous Galerkin method for the numerical solution of inviscid compressible flows. *Journal of Computational Physisc*, 198:727–746, 2004.

- [52] V. Dolejsi, M. Feistauer, and J. Hozman. Analysis of semi-implicit DGFEM for nonlinear convection-diffusion problems on nonconforming meshes. *Computer Methods in Applied Mechanics and Engineering*, 196:2813–2827, 2007.
- [53] M. Dumbser. *Arbitrary High Order Schemes for the Solution of Hyperbolic Conservation Laws in Complex Domains*. Shaker Verlag, 2005.
- [54] M. Dumbser and V. Casulli. A staggered semi-implicit spectral discontinuous Galerkin scheme for the shallow water equations. *Applied Mathematics and Computation*, 219(15):8057–8077, 2013.
- [55] M. Dumbser and V. Casulli. A conservative, weakly nonlinear semi-implicit finite volume scheme for the compressible Navier-Stokes equations with general equation of state. *Applied Mathematics and Computation*, 272:479–497, 2016.
- [56] M. Dumbser, U. Iben, and M. Ioriatti. An efficient semi-implicit finite volume method for axially symmetric compressible flows in compliant tubes. *Applied Numerical Mathematics*, 89:24–44, 2015.
- [57] M. Dumbser and M. Käser. Arbitrary High Order Non-oscillatory Finite Volume Schemes on Unstructured Meshes for Linear Hyperbolic Systems. *Journal of Computational Physics*, 221(2):693–723, February 2007.
- [58] M. Dumbser, M. Käser, and E. F. Toro. An Arbitrary High Order Discontinuous Galerkin Method for Elastic Waves on Unstructured Meshes V: Local Time Stepping and  $p$ -Adaptivity. *Geophysical Journal International*, 171:695–717, 2007.
- [59] M. Dumbser and R. Loubère. A simple robust and accurate a posteriori sub-cell finite volume limiter for the discontinuous Galerkin method on unstructured meshes. *Journal of Computational Physics*, 319(Supplement C):163 – 199, 2016.
- [60] M. Dumbser, I. Peshkov, E. Romenski, and O. Zanotti. High order ADER schemes for a unified first order hyperbolic formulation of continuum mechanics: Viscous heat-conducting fluids and elastic solids. *Journal of Computational Physics*, 314:824 – 862, 2016.
- [61] M. Dumbser, I. Peshkov, E. Romenski, and O. Zanotti. High order ADER schemes for a unified first order hyperbolic formulation of Newtonian continuum mechanics coupled with electro-dynamics. *Journal of Computational Physics*, 348:298 – 342, 2017.
- [62] M. Dumbser and E. F. Toro. A simple extension of the Osher Riemann solver to non-conservative hyperbolic systems. *Journal of Scientific Computing*, 48:70–88, 2011.
- [63] M. Dumbser and E.F. Toro. A Simple Extension of the Osher Riemann Solver to Non-conservative Hyperbolic Systems. *Journal of Scientific Computing*, 48(1):70–88, Jul 2011.
- [64] M. Dumbser, O. Zanotti, R. Loubère, and S. Diot. A posteriori subcell limiting of the discontinuous Galerkin finite element method for hyperbolic conservation laws. *Journal of Computational Physics*, 278:47 – 75, 2014.

- [65] Michael Dumbser and Dinshaw S. Balsara. A new efficient formulation of the hllem riemann solver for general conservative and non-conservative hyperbolic systems. *Journal of Computational Physics*, 304:275 – 319, 2016.
- [66] Michael Dumbser, Dinshaw S. Balsara, Maurizio Tavelli, and Francesco Fambri. A divergence-free semi-implicit finite volume scheme for ideal, viscous and resistive magnetohydrodynamics. *International Journal for Numerical Methods in Fluids*.
- [67] Michael Dumbser and Eleuterio F. Toro. On universal osher-type schemes for general nonlinear hyperbolic conservation laws. *Communications in Computational Physics*, 10(3):635–671, 2011.
- [68] B. Einfeldt. On godunov-type methods for gas dynamics. *SIAM Journal of Numerical Analysis*, 25:294 – 318, 1988.
- [69] B. Einfeldt, C. D. Munz, P. L. Roe, and B. Sjögreen. On godunov-type methods near low densities. *Journal of Computational Physics*, 92:273–295, 1991.
- [70] F. Fambri and M. Dumbser. Spectral semi-implicit and space-time discontinuous Galerkin methods for the incompressible Navier-Stokes equations on staggered Cartesian grids. *Applied Numerical Mathematics*, 110:41–74, 2016.
- [71] F. Fambri and M. Dumbser. Semi-implicit discontinuous Galerkin methods for the incompressible Navier-Stokes equations on adaptive staggered Cartesian grids. *Computer Methods in Applied Mechanics and Engineering*, 324:170 – 203, 2017.
- [72] F. Fambri, M. Dumbser, and V. Casulli. An Efficient Semi-Implicit Method for Three-Dimensional Non-Hydrostatic Flows in Compliant Arterial Vessels. *International Journal for Numerical Methods in Biomedical Engineering*, 30:1170–1198, 2014.
- [73] Francesco Fambri, Michael Dumbser, Sven Koppel, Giuliano Rezzolla, and Olindo Zanotti. Ader discontinuous galerkin schemes for general-relativistic ideal magnetohydrodynamics. *Monthly Notices of the Royal Astronomical Society*, 477:4543 – 4564, 2018.
- [74] Francesco Fambri, Michael Dumbser, and Olindo Zanotti. Space-time adaptive ader-dg schemes for dissipative flows: Compressible navier-stokes and resistive mhd equations. *Computer Physics Communications*, 220:297 – 318, 2017.
- [75] L. Formaggia and A. Scotti. Positivity and Conservation Properties of Some Integration Schemes for Mass Action Kinetics. *SIAM Journal on Numerical Analysis*, 49(3):1267–1288, 2011.
- [76] Elena Gaburro, Michael Dumbser, and Manuel J. Castro. Direct arbitrary-lagrangian-eulerian finite volume schemes on moving nonconforming unstructured meshes. *Computers and Fluids*, 159:254 – 275, 2017.
- [77] Elena Gaburro, Michael Dumbser, and Manuel J. Castro. Well-balanced arbitrary-lagrangian-eulerian finite volume schemes on moving nonconforming meshes for the euler equations of gas dynamics with gravity. *Monthly Notices of the Royal Astronomical Society*, 477:2251 – 2275, 2018.



- 
- [78] S.K. Godunov. Finite difference methods for the computation of discontinuous solutions of the equations of fluid dynamics. *Math. USSR Sb.*, 47:271–306, 1959.
- [79] M. Goeke. *Modellierung der Materialdämpfung bei der Wellenausbreitung in flexiblen Leitungen*. PhD thesis, Institut für Baustatik und Baudynamik der Universität Stuttgart, 2013.
- [80] L. Gosse. A well-balanced flux-vector splitting scheme designed for hyperbolic systems of conservation laws with source terms. *Computers and Mathematics with Applications*, 39(9):135 – 159, 2000.
- [81] Laurent Gosse. A well-balanced scheme using non-conservative products designed for hyperbolic systems of conservation laws with source terms. *Mathematical Models and Methods in Applied Sciences*, 11(02):339–365, 2001.
- [82] Sigal Gottlieb and Chi-Wang Shu. Total Variation Diminishing Runge-Kutta Schemes. *Mathematics of Computation*, 67(221):73–85, January 1998.
- [83] J. M. Greenberg and A. Y. Leroux. A Well-Balanced Scheme for the Numerical Processing of Source Terms in Hyperbolic Equations. *SIAM Journal on Numerical Analysis*, 33(1):1–16, 1996.
- [84] F.H. Harlow and J.E. Welch. Numerical calculation of time-dependent viscous incompressible flow of fluid with a free surface. *Physics of Fluids*, 8:2182–2189, 1965.
- [85] A. Harten, P.D. Lax, and B. van Leer. On upstream differencing and godunov-type schemes for hyperbolic conservation laws. *SIAM Review*, 25(1):35–61, 1983.
- [86] Ami Harten, Bjorn Engquist, Stanley Osher, and Sukumar R. Chakravarthy. Uniformly high order accurate essentially non-oscillatory schemes, iii. *Journal of Computational Physics*, 131(1):3 – 47, 1997.
- [87] M. R. Hestenes and E. Stiefel. Methods of conjugate gradients for solving linear systems. *Journal of Research of the National Bureau of Standards*, 49:409–436, 1952.
- [88] Arturo Hidalgo and Michael Dumbser. Ader schemes for nonlinear systems of stiff advection–diffusion–reaction equations. *Journal of Scientific Computing*, 48(1):173–189, Jul 2011.
- [89] Rouleau WT Holmboe EL. The effect of viscous shear on transients in liquid lines. *ASME Journal of Basic Engineering*, 89:174–180, 1967.
- [90] M. Ioriatti and M. Dumbser. A posteriori sub-cell finite volume limiting of staggered semi-implicit discontinuous galerkin schemes for the shallow water equations. *Applied Numerical Mathematics*, 2018. to appear.
- [91] M. Ioriatti, M. Dumbser, and U. Iben. A comparison of explicit and semi-implicit finite volume schemes for viscous compressible flows in elastic pipes in fast transient regime. *Zeitschrift fuer Angewandte Mathematik und Mechanik*, 97:1358–1380, 2017.

- [92] Matteo Ioriatti and Michael Dumbser. Semi-implicit staggered discontinuous galerkin schemes for axially symmetric viscous compressible flows in elastic tubes. *Computers & Fluids*, 167:166 – 179, 2018.
- [93] Guang-Shan Jiang and Chi-Wang Shu. Efficient implementation of weighted eno schemes. *Journal of Computational Physics*, 126(1):202 – 228, 1996.
- [94] J.Qiu and C-W.Shu. Hermite WENO Schemes and Their Application As Limiters for Runge-Kutta Discontinuous Galerkin Method: One-dimensional Case. *Journal of Computational Physics*, 193(1):115–135, January 2004.
- [95] C.W. Shu J.Zhu, X.Zhong and J. Qiu. Runge-Kutta discontinuous Galerkin method using a new type of WENO limiters on unstructured meshes. *Journal of Computational Physics*, 248:200–220, 2013.
- [96] Toshiharu Kagawa, I Lee, A Kitagawa, and T Takenaka. High speed and accurate computing method of frequency dependent friction in laminar pipe flow fow characteristics method. 49:2638–2644, 01 1983.
- [97] M. Käser and A. Iske. ADER Schemes on Adaptive Triangular Meshes for Scalar Conservation Laws. *Journal of Computational Physics*, 205(2):486–508, May 2005.
- [98] C.M. Klaij, J.J.W. van der Vegt, and H. van der Ven. Space-time discontinuous Galerkin method for the compressible Navier-Stokes equations. *Journal of Computational Physics*, 217(2):589 – 611, 2006.
- [99] S. C. Kramer and G. S. Stelling. A conservative unstructured scheme for rapidly varied flows. *International Journal for Numerical Methods in Fluids*, 58:183–212, 2008.
- [100] L. Krivodonova. Limiters for high–order discontinuous Galerkin methods. *Journal of Computational Physics*, 226:879–896, 2007.
- [101] L. Krivodonova, J. Xin, J.F. Remacle, N. Chevaugeon, and J.E. Flaherty. Shock detection and limiting with discontinuous Galerkin methods for hyperbolic conservation laws. *Applied Numerical Mathematics*, 48:323–338, 2004.
- [102] Alexander Kurganov and Eitan Tadmor. Solution of two-dimensional riemann problems for gas dynamics without riemann problem solvers. *Numerical Methods for Partial Differential Equations*, 18(5):584–608.
- [103] Peter Lax and Burton Wendroff. Systems of conservation laws. *Communications on Pure and Applied Mathematics*, 13(2):217–237.
- [104] J. Leibinger, M. Dumbser, U. Iben, and I. Wayand. A path-conservative Osher-type scheme for axially symmetric compressible flows in flexible visco-elastic tubes. *Applied Numerical Mathematics*, 105:47–63, 2016.
- [105] R. Leonhardt. *Dynamische Untersuchungen von Hydraulikkomponenten*. PhD thesis, University of Karlsruhe, 2008.

- [106] R.J. LeVeque. Balancing source terms and flux gradients in high-resolution Godunov methods: The quasi-steady wave-propagation algorithm. *Journal of Computational Physics*, 146(1), Oct 1998.
- [107] R. Loubère, M. Dumbser, and S. Diot. A New Family of High Order Unstructured MOOD and ADER Finite Volume Schemes for Multidimensional Systems of Hyperbolic Conservation Laws. *Communications in Computational Physics*, 16(3):718–763, 2014.
- [108] M. Yang and Z. Wang. A Parameter-Free Generalized Moment Limiter for High-Order Methods on Unstructured Grids. *Advances in Applied Mathematics and Mechanics*, 2009.
- [109] Colin B. Macdonald, Sigal Gottlieb, and Steven J. Ruuth. A Numerical Study of Diagonally Split Runge–Kutta Methods for PDEs with Discontinuities. *Journal of Scientific Computing*, 36(1):89–112, Jul 2008.
- [110] F.C. Massa, G. Noventa, M. Lorini, F. Bassi, and A. Ghidoni. High-order linearly implicit two-step peer schemes for the discontinuous Galerkin solution of the incompressible Navier-Stokes equations. *Computers and Fluids*, 162:55–71, 2018.
- [111] Andreas Meister and Sigrun Ortleb. A positivity preserving and well-balanced dg scheme using finite volume subcells in almost dry regions. *Applied Mathematics and Computation*, 272:259 – 273, 2016. Recent Advances in Numerical Methods for Hyperbolic Partial Differential Equations.
- [112] Peter Moczo and Jozef Kristek. On the rheological models used for time-domain methods of seismic wave propagation. *Geophysical Research Letters*, 32(1).
- [113] Sudi Mungkasi and Stephen G. Roberts. On the best quantity reconstructions for a well balanced finite volume method used to solve the shallow water wave equations with a wet/dry interface. In P. Howlett, M. Nelson, and A. J. Roberts, editors, *Proceedings of the 9th Biennial Engineering Mathematics and Applications Conference, EMAC-2009*, volume 51 of *ANZIAM J.*, pages C48–C65, March 2010. <http://anziamj.austms.org.au/ojs/index.php/ANZIAMJ/article/view/2576> [March 3, 2010].
- [114] J. Von Neumann and R. D. Richtmyer. A method for the numerical calculation of hydrodynamic shocks. *Journal of Applied Physics*, 21:232–237, 1950.
- [115] S. Osher. Riemann solvers, the entropy condition, and difference. *SIAM Journal on Numerical Analysis*, 21(2):217–235, 1984.
- [116] S. Osher and F. Solomon. Upwind difference schemes for hyperbolic conservation laws. *Mathematics of Computation*, 38:339–374, 1982.
- [117] C. Parés. Numerical methods for nonconservative hyperbolic systems: a theoretical framework. *SIAM Journal on Numerical Analysis*, 44(1):300–321, 2006.
- [118] P.-O. Persson and J. Peraire. Sub-cell shock capturing for discontinuous Galerkin methods. *AIAA Paper 2006-112*, 2006.

- [119] Peshkov, I. and Romenski, E. A hyperbolic model for viscous Newtonian flows. *Continuum Mechanics and Thermodynamics*, 28:85–104, 2016.
- [120] Giuseppe Pezzinga. Evaluation of unsteady flow resistances by quasi-2d or 1d models. *Journal of Hydraulic Engineering*, 126(10):778–785, 2000.
- [121] L. Prandtl. Über Flüssigkeitsbewegung bei sehr kleiner Reibung. *Verhandlg. III. Intern. Math. Kongr. Heidelberg*, pages 484–491, 1904.
- [122] Jianxian Qiu, Michael Dumbser, and Chi-Wang Shu. The discontinuous galerkin method with lax-wendroff type time discretizations. *Computer Methods in Applied Mechanics and Engineering*, 194(42):4528 – 4543, 2005.
- [123] Alfio Quarteroni, Riccardo Sacco, and Fausto Saleri. *Numerical Mathematic*. Springer, 2007.
- [124] W. H. Reed and T. R. Hill. Triangular mesh methods for the neutron transport equation. Technical Report LA-UR-73-479, Los Alamos Scientific Laboratory, 1973.
- [125] P.L. Roe. Approximate Riemann solvers, parameter vectors, and difference schemes. *Journal of Computational Physics*, 43:357–372, 1981.
- [126] V. V. Rusanov. Calculation of interaction of non-steady shock waves with obstacles. *J. Comput. Math. Phys. USSR*, 1:267 – 305, 1961.
- [127] Y. Saad and M. Schultz. Gmres: A generalized minimal residual algorithm for solving non-symmetric linear systems. *SIAM Journal on Scientific and Statistical Computing*, 7(3):856–869, 1986.
- [128] H. Schlichting and K. Gersten. *Grenzschichttheorie*. Springer Verlag, 2005.
- [129] C. Schulz-Rinne. Classification of the riemann problem for two-dimensional gas dynamics. *SIAM Journal on Mathematical Analysis*, 24(1):76–88, 1993.
- [130] Chi-Wang Shu. *Essentially non-oscillatory and weighted essentially non-oscillatory schemes for hyperbolic conservation laws*, pages 325–432. Springer Berlin Heidelberg, Berlin, Heidelberg, 1998.
- [131] Gary A Sod. A survey of several finite difference methods for systems of nonlinear hyperbolic conservation laws. *Journal of Computational Physics*, 27(1):1 – 31, 1978.
- [132] M. Sonntag and C.D. Munz. Shock Capturing for discontinuous Galerkin methods using Finite Volume Subcells. In J. Fuhrmann, M. Ohlberger, and C. Rohde, editors, *Finite Volumes for Complex Applications VII*, pages 945–953. Springer, 2014.
- [133] M. Sonntag and C.D. Munz. Efficient Parallelization of a Shock Capturing for Discontinuous Galerkin Methods using Finite Volume Sub-cells. *Journal of Scientific Computing*, 70:1262–1289, 2017.

- [134] G. S. Stelling and S. P. A. Duynmeyer. A staggered conservative scheme for every Froude number in rapidly varied shallow water flows. *International Journal for Numerical Methods in Fluids*, 43:1329–1354, 2003.
- [135] Arne Taube, Michael Dumbser, Dinshaw S. Balsara, and Claus-Dieter Munz. Arbitrary high-order discontinuous galerkin schemes for the magnetohydrodynamic equations. *Journal of Scientific Computing*, 30(3):441–464, Mar 2007.
- [136] M. Tavelli and M. Dumbser. A high order semi-implicit discontinuous Galerkin method for the two dimensional shallow water equations on staggered unstructured meshes. *Applied Mathematics and Computation*, 234:623–644, 2014.
- [137] M. Tavelli and M. Dumbser. A staggered semi-implicit discontinuous Galerkin method for the two dimensional incompressible Navier-Stokes equations. *Applied Mathematics and Computation*, 248:70–92, 2014.
- [138] M. Tavelli and M. Dumbser. A staggered space-time discontinuous Galerkin method for the incompressible Navier-Stokes equations on two-dimensional. *Computers and Fluids*, 119:235–249, 2015.
- [139] M. Tavelli and M. Dumbser. A staggered space-time discontinuous Galerkin method for the three-dimensional incompressible Navier-Stokes equations on unstructured tetrahedral meshes. *Journal of Computational Physics*, 319:294–323, 2016.
- [140] M. Tavelli and M. Dumbser. A pressure-based semi-implicit space-time discontinuous Galerkin method on staggered unstructured meshes for the solution of the compressible Navier-Stokes equations at all Mach numbers. *Journal of Computational Physics*, 341:341 – 376, 2017.
- [141] M. Tavelli, M. Dumbser, and V. Casulli. High resolution methods for scalar transport problems in compliant systems of arteries. *Applied Numerical Mathematics*, 74:62–82, 2013.
- [142] Maurizio Tavelli and Michael Dumbser. Arbitrary high order accurate space-time discontinuous galerkin finite element schemes on staggered unstructured meshes for linear elasticity. *Journal of Computational Physics*, 366:386 – 414, 2018.
- [143] Maurizio Tavelli, Michael Dumbser, Dominic Etienne Charrier, Leonhard Rannabauer, Tobias Weinzierl, and Michael Bader. A simple diffuse interface approach on adaptive cartesian grids for the linear elastic wave equations with complex topography. *Journal of Computational Physics*.
- [144] William Carlisle Thacker. Some exact solutions to the nonlinear shallow-water wave equations. *Journal of Fluid Mechanics*, 107:499–508, 1981.
- [145] V. A. Titarev and E. F. Toro. Ader: Arbitrary high order godunov approach. *Journal of Scientific Computing*, 17(1):609–618, Dec 2002.
- [146] Vladimir A. Titarev and Eleuterio. F. Toro. Ader schemes for three-dimensional nonlinear hyperbolic systems. 2004.

- [147] E. F. Toro. *Shock-Capturing Methods for Free-Surface Shallow Flows*. Wiley, 2001.
- [148] E. F. Toro. *Riemann Solvers and Numerical Methods for Fluid Dynamics*. Springer, 2009.
- [149] E.F. Toro, M. Spruce, and W. Speares. Restoration of the contact surface in the harten-lax-van leer riemann solver. *Journal of Shock Waves*, 4:25–34, 1994.
- [150] E.F. Toro and V.A. Titarev. Ader schemes for scalar non-linear hyperbolic conservation laws with source terms in three-space dimensions. *Journal of Computational Physics*, 202(1):196 – 215, 2005.
- [151] E.F. Toro and V.A. Titarev. Derivative riemann solvers for systems of conservation laws and ader methods. *Journal of Computational Physics*, 212(1):150 – 165, 2006.
- [152] E.F. Toro and M.E. Vázquez-Cendón. Flux splitting schemes for the Euler equations. *Computers and Fluids*, 70:1 – 12, 2012.
- [153] A. K. Trikha. An efficient method for simulating frequency-dependent friction in transient liquid flow. *ASME Journal of Basic Engineering*, 97:97–105, 1975.
- [154] G. Tumolo, L. Bonaventura, and M. Restelli. A semi-implicit, semi-Lagrangian, p-adaptive discontinuous Galerkin method for the shallow water equations. *Journal of Computational Physics*, 232:46–67, 2013.
- [155] Giovanni Tumolo and Luca Bonaventura. A semi-implicit, semi-Lagrangian discontinuous Galerkin framework for adaptive numerical weather prediction. *Quarterly Journal of the Royal Meteorological Society*, 141(692):2582–2601.
- [156] J.J.W. van der Vegt and H. van der Ven. Space-Time Discontinuous Galerkin Finite Element Method with Dynamic Grid Motion for Inviscid Compressible Flows: I. General Formulation. *Journal of Computational Physics*, 182(2):546 – 585, 2002.
- [157] B. van Leer. Toward the ultimate conservative difference scheme. v. a second-order sequel to godunov’s method. *Journal of Computational Physics*, 32:101 – 136, 1979.
- [158] Alan Vardy and Jim Brown. On turbulent, unsteady, smooth-pipe friction. *Pressure Surges and Fluid Transients*, BHR Group, pages 289–311, 1996.
- [159] Alan Vardy and Jim Brown. Efficient approximation of unsteady friction weighting functions. *Journal of Hydraulic Engineering*, 130(11):1097–1107, 2004.
- [160] John Vítkovský, Mark Stephens, Anton Bergant, Martin Lambert, Angus Simpson, and A Simpson. Efficient and accurate calculation of zielke and varyd-brown unsteady friction in pipe transients. 03 2004.
- [161] J. R. Womersley. Method for the calculation of velocity, rate of flow and viscous drag in arteries when the pressure gradient is known. *J. Physiol.*, 127:553–563, 1955.

- [162] W.Zielke. Frequency-dependent friction in transient pipe flow. *Journal of Basic Engineering*, 99:109–115, 1968.
- [163] O. Zanotti, F. Fambri, M. Dumbser, and A. Hidalgo. Space-time adaptive ADER discontinuous Galerkin finite element schemes with a posteriori sub-cell finite volume limiting. *Computer and Fluids*, 118:204–224, 2015.
- [164] Olindo Zanotti, Francesco Fambri, and Michael Dumbser. Solving the relativistic magnetohydrodynamics equations with ader discontinuous galerkin methods, a posteriori subcell limiting and adaptive mesh refinement. *Monthly Notices of the Royal Astronomical Society*, 452:3010 – 3029, 2015.
- [165] K.Urbanowicz Z. Zarzycki. New efficient approximation of weighting functions for simulations of unsteady losses in liquid pipe flow. *Journal of Theoretical and Applied Mechanics*, 20:487–508, 2012.





# List of Tables

2.1	Frequency domain - error $\epsilon_{ z_{11} }$ for the first order schemes and for the second order schemes corresponding to different $N_x$ . . . . .	55
2.2	Frequency domain - CPU times for the first order schemes and for the second order schemes corresponding to different $N_x$ . . . . .	56
2.3	CPU times in seconds for the two-dimensional method compared with different one dimensional methods . . . . .	57
3.1	Numerical convergence rates computed with respect to the $L_2$ error norms of pressure and velocity for the semi-implicit staggered DG scheme applied to the 1D model.	78
3.2	Numerical convergence rates computed with respect to the $L_2$ error norms of pressure and velocity for the staggered semi-implicit DG scheme applied to the $2D_{x^r}$ model. . . . .	79
4.1	Numerical validation of the well-balanced property for the semi-implicit staggered DG scheme in 1D with sub-cell FV limiter. $L_\infty$ and $L_2$ norms for $\eta$ with $P = 5$ at time $t = 0.1$ for different machine precisions. . . . .	124
4.2	Riemann problems for semi-implicit DG on staggered grid with sub-cell limiter - Left and right states for $\eta, u$ and $h$ , left and right boundaries $x_L$ and $x_R$ , number of control volume $N_x$ , $\Delta_x$ and final time of the simulations $t_{end}$ . . . . .	131
4.3	Numerical validation of the well-balanced property for the semi-implicit staggered DG scheme with subcell FV limiter in 2D. $L_\infty$ and $L_2$ norms for $\eta$ with $P = 5$ at time $t = 0.1$ for different machine precisions. . . . .	136
5.1	Riemann problems for semi-implicit DG on staggered grid with sub-cell limiter - Left and right states for $\rho, U$ and $p$ , $x_0$ , number of control volume $N_x$ , $\Delta_x$ and final time of the simulations $t_{end}$ . . . . .	165
5.2	Shu-Osher vortex - Numerical convergence rates computed with respect to the $L_2$ error norms of pressure and velocity for the two-dimensional semi-implicit staggered DG scheme. . . . .	174
5.3	Initial conditions for the two-dimensional Riemann problems. . . . .	182
5.4	Initial condition for the 2D Riemann problem . . . . .	182



# List of Figures

2.1	Grid for the semi-implicit 1D scheme: according to the staggered approach, the pressure is defined in the cell barycenter while the velocity is defined in the cell edges	16
2.2	Radial grid for the axial velocity $u$	19
2.3	Radial grid for the radial velocity $w$	20
2.4	Womersley test, $\mu = 10^{-3}$ . Exact solution (—), numerical data given by SI2D (▲) and by SI1D with Zielke model (▽)	25
2.5	Womersley test, $\mu = 10^{-3}$ . Exact solution (—), numerical data given by SI1D with Trikha model (▲) and with Brunone model (▽)	25
2.6	Womersley test, $\mu = 10^{-3}$ . Exact solution (—), numerical data given by SI1D with Kagawa model (▲) and with Kagawa ODE model (▽)	26
2.7	Womersley test, $\mu = 10^{-3}$ . Exact solution (—), numerical data given by SI1D with UZ model (▲) and with UZ ODE model (▽)	26
2.8	Womersley test, $\mu = 1$ . Exact solution (—), numerical data given by SI2D (▲) and by SI1D with Zielke model (▽)	27
2.9	Womersley test, $\mu = 1$ . Exact solution (—), numerical data given by SI1D with Trikha model (▲) and with Brunone model (▽)	27
2.10	Womersley test, $\mu = 1$ . Exact solution (—), numerical data given by SI1D with Kagawa model (▲) and with Kagawa ODE model (▽)	28
2.11	Womersley test, $\mu = 1$ . Exact solution (—), numerical data given by SI1D with UZ model (▲) and with UZ ODE model (▽)	28
2.12	Absolute value and phase of the element $z_{11}$ of the impedance matrix for a test fluid in a rigid pipe - Reference solution (—), experimental data (●) and numerical data from the method SI2D ▲	31
2.13	Absolute value and phase of the element $z_{11}$ of the impedance matrix for a test fluid in a rigid pipe - Reference solution (—), experimental data (●) and numerical data from the method SI1D (▲) and from PC1D (▽) using only a steady friction model	32
2.14	Absolute value and phase of the element $z_{11}$ of the impedance matrix for a test fluid in a rigid pipe - Reference solution (—), experimental data (●) and numerical data from the method SI1D (▲) and from PC1D (▽) with Trikha model for the unsteady friction term	33
2.15	Absolute value and phase of the element $z_{11}$ of the impedance matrix for a test fluid in a rigid pipe - Reference solution (—), experimental data (●) and numerical data from the method SI1D (▲) and from PC1D (▽) with Kagawa model for the unsteady friction term	34
2.16	Absolute value and phase of the element $z_{11}$ of the impedance matrix for a test fluid in a rigid pipe - Reference solution (—), experimental data (●) and numerical data from the method SI1D (▲) and from PC1D (▽) with Kagawa ODE model for the unsteady friction term	35

2.17	Absolute value and phase of the element $z_{11}$ of the impedance matrix for a test fluid in a rigid pipe - Reference solution (—), experimental data (●) and numerical data from the method SI1D (▲) and from PC1D (▽) with Urbanowicz and Zarzycky model for the unsteady friction term . . . . .	36
2.18	Absolute value and phase of the element $z_{11}$ of the impedance matrix for a test fluid in a rigid pipe - Reference solution (—), experimental data (●) and numerical data from the method SI1D (▲) and from PC1D (▽) with Urbanowicz and Zarzycky ODE model for the unsteady friction term . . . . .	37
2.19	Absolute value and phase of the element $z_{11}$ of the impedance matrix for a test fluid in a rigid pipe - Reference solution (—), experimental data (●) and numerical data from the method SI1D (▲) and from PC1D (▽) with Brunone model for the unsteady friction term . . . . .	38
2.20	Absolute value and phase of the element $z_{11}$ of the impedance matrix for a compressible fluid in an elastic pipe - Reference solution (—) and numerical data from the method SI2D ▲ . . . . .	40
2.21	Absolute value and phase of the element $z_{11}$ of the impedance matrix for a compressible fluid in an elastic pipe - Reference solution (—) and numerical data from the method SI1D (▲) and from PC1D (▽) accounting only the steady friction term . . . . .	41
2.22	Absolute value and phase of the element $z_{11}$ of the impedance matrix for a compressible fluid in an elastic pipe - Reference solution (—) and numerical data from the method SI1D (▲) and from PC1D (▽) with Trikha model for the unsteady friction term . . . . .	42
2.23	Absolute value and phase of the element $z_{11}$ of the impedance matrix for a compressible fluid in an elastic pipe - Reference solution (—) and numerical data from the method SI1D (▲) and from PC1D (▽) with Kagawa model for the unsteady friction term . . . . .	43
2.24	Absolute value and phase of the element $z_{11}$ of the impedance matrix for a compressible fluid in an elastic pipe - Reference solution (—) and numerical data from the method SI1D (▲) and from PC1D (▽) with Kagawa ODE model for the unsteady friction term . . . . .	44
2.25	Absolute value and phase of the element $z_{11}$ of the impedance matrix for a compressible fluid in an elastic pipe - Reference solution (—) and numerical data from the method SI1D (▲) and from PC1D (▽) with Urbanowicz and Zarzycky model for the unsteady friction term . . . . .	45
2.26	Absolute value and phase of the element $z_{11}$ of the impedance matrix for a compressible fluid in an elastic pipe - Reference solution (—) and numerical data from the method SI1D (▲) and from PC1D (▽) with Urbanowicz and Zarzycky ODE model for the unsteady friction term . . . . .	46
2.27	Absolute value and phase of the element $z_{11}$ of the impedance matrix for a compressible fluid in an elastic pipe - Reference solution (—) and numerical data from the method SI1D (▲) and from PC1D (▽) with Brunone model for the unsteady friction term . . . . .	47
2.28	Water hammer, rigid pipe - Reference solution (—), 1D methods without unsteady friction (- - -), numerical data given by SI1D with Trikha model (▲) and by PC1D with Trikha model (▽) . . . . .	48

2.29	Water hammer, rigid pipe - Reference solution (—), 1D methods without unsteady friction (---), numerical data given by SI1D with Brunone model (▲) and by PC1D with Brunone model (▽) . . . . .	48
2.30	Water hammer, rigid pipe - Reference solution (—), 1D methods without unsteady friction (---), numerical data given by SI1D with Kagawa model (▲) and by PC1D with Kagawa model (▽) . . . . .	49
2.31	Water hammer, rigid pipe - Reference solution (—), 1D methods without unsteady friction (---), numerical data given by SI1D with Kagawa ODE model (▲) and by PC1D with Kagawa ODE model (▽) . . . . .	49
2.32	Water hammer, rigid pipe - Reference solution (—), 1D methods without unsteady friction (---), numerical data given by SI1D with UZ model (▲) and by PC1D with UZ model (▽) . . . . .	50
2.33	Water hammer, rigid pipe - Reference solution (—), 1D methods without unsteady friction (---), numerical data given by SI1D with UZ ODE model (▲) and by PC1D with UZ ODE model (▽) . . . . .	50
2.34	Water hammer, elastic pipe - Reference solution (—), 1D methods without unsteady friction (---), numerical data given by SI1D with Kagawa ODE model (▲) and by PC1D with Kagawa ODE model (▽) . . . . .	51
2.35	Water hammer, elastic pipe - Reference solution (—), 1D methods without unsteady friction (---), numerical data given by SI1D with UZ ODE model (▲) and by PC1D with UZ ODE model (▽) . . . . .	51
2.36	$N_x$ and errors for the first order schemes - ● Explicit schemes, ▲ SI1D, ■ SI2D .	53
2.37	$N_x$ and errors for the second order schemes - ● Explicit schemes, ▲ SI1D, ■ SI2D	53
2.38	CPU times and errors for the first order schemes - ● Explicit schemes, ▲ SI1D, ■ SI2D . . . . .	54
2.39	CPU times and errors for the second order schemes - ● Explicit schemes, ▲ SI1D, ■ SI2D . . . . .	54
3.1	Staggered control volumes for the DG scheme for the 1D model. Main grid used for the pressure (top) and staggered dual mesh for the velocity (bottom). . . . .	62
3.2	Steady flow in an elastic pipe. From top left to bottom right: exact solution (—) and numerical data for pressure, radius and mean velocity computed using the semi-implicit staggered DG scheme for the 1D model (□) and for the 2D <sub>xr</sub> model (◇). The radial velocity profiles in the last figure can be computed only using the 2D <sub>xr</sub> model and are shown for the inlet of the pipe (▲) and the outlet of the pipe (▼). . . . .	77
3.3	Womersley profiles for Test A (left) and Test B (right). Analytical solution (—) and numerical data from the 2D <sub>xr</sub> model after 25 (▽), 50 (◇) and 75 (△) timesteps for the axial velocity $u$ . . . . .	81
3.4	Absolute value and phase of the elements $z_{11}$ and $z_{12}$ of the impedance matrix for a fluid in a rigid pipe - Reference solution (—) and numerical data using the semi-implicit staggered DG scheme applied to the 1D model (□) and to the 2D <sub>xr</sub> model (◇). . . . .	83
3.5	Absolute value and phase of the elements $z_{11}$ and $z_{12}$ of the impedance matrix for a fluid in an elastic pipe - Reference solution (—) and numerical data using the semi-implicit staggered DG scheme applied to the 1D model (□) and to the 2D <sub>xr</sub> model (◇). . . . .	84

3.6	Blasius boundary layer. Reference solution (—) and numerical results using the staggered SIDG scheme applied to the $2D_{xT}$ model ( $\square$ ). . . . .	85
4.1	Staggered grids for the one-dimensional DG scheme. Main grid used for the free surface (top) and dual mesh for the velocity (bottom). . . . .	89
4.2	Staggered grids for the sub-cell finite volume scheme in one dimension in the case $P = 2$ ( $N_s = 5$ ). Main grid used for the free surface (top) and staggered dual mesh for the velocity (bottom). . . . .	93
4.3	Projection and reconstruction operators - Operators $\mathcal{P}$ and $\mathcal{W}$ . . . . .	97
4.4	Projection and reconstruction operators - Operator $\mathcal{W}_L$ . . . . .	97
4.5	Projection and reconstruction operators - Operator $\mathcal{W}_R$ . . . . .	98
4.6	Staggered grids for the semi-implicit staggered DG scheme with sub-cell limiter active in cell $T_i$ . Main grid used for the free surface (top) and staggered dual mesh for the velocity (bottom). If a cell $T_i$ on the main grid is flagged as troubled, then also the two overlapping staggered velocity control volumes $T_{i\pm\frac{1}{2}}$ are flagged as troubled. The data representation in troubled cells is changed from high order polynomials to piecewise constant subcell averages. . . . .	99
4.7	Control volumes for the continuity equation for the staggered semi-implicit DG scheme in the limited case using $P = 2$ , i.e. $N_s = 2P + 1 = 5$ . DG method used in $T_i$ and $T_{i-\frac{1}{2}}$ ; finite volumes in $T_{i+\frac{1}{2}}$ . . . . .	100
4.8	Control volumes for the continuity equation for the staggered semi-implicit DG scheme in the limited case using $P = 2$ , i.e. $N_s = 2P + 1 = 5$ . DG method used in $T_i$ and $T_{i+\frac{1}{2}}$ ; finite volumes in $T_{i-\frac{1}{2}}$ . . . . .	101
4.9	Control volumes for the momentum equation for the staggered semi-implicit DG scheme in the limited case using $P = 2$ ( $N_s = 5$ ). DG in $T_i$ ; FV in $T_{i+1}$ and in $T_{i+\frac{1}{2}}$ . . . . .	102
4.10	Control volumes for the momentum equation for the staggered semi-implicit DG scheme in the limited case using $P = 2$ ( $N_s = 5$ ). DG in $T_{i+1}$ ; FV in $T_i$ and in $T_{i+\frac{1}{2}}$ . . . . .	103
4.11	Computational grids for the 2D semi-implicit staggered DG scheme. Main grid for $\eta$	107
4.12	Computational grids for the 2D semi-implicit staggered DG scheme. Staggered dual meshes for $U$ and $V$ . . . . .	108
4.13	Main grid and stencil for the 2D semi-implicit sub-cell finite volume scheme. . . . .	112
4.14	Ordering of the subcell degrees of freedom within one cell for the 2D semi-implicit sub-cell finite volume scheme . . . . .	113
4.15	Staggered mesh in $x$ direction for the 2D semi-implicit sub-cell finite volume scheme	114
4.16	Staggered mesh in $y$ direction for the 2D semi-implicit sub-cell finite volume scheme	115
4.17	Subcell finite volumes in two space dimensions. Reference domains $\Omega_L^x$ and $\Omega_R^x$ for the reconstructions . . . . .	116
4.18	Subcell finite volumes in two space dimensions. Reference domains $\Omega_L^y$ and $\Omega_R^y$ for the reconstructions . . . . .	117
4.19	Reference solutions and numerical solutions for the LeVeque test with $\epsilon = 0.2$ at $t = 0.2$ . . . . .	125
4.20	Reference solutions and numerical solutions for the LeVeque test with $\epsilon = 10^{-3}$ at $t = 0.2$ . . . . .	125

4.21	Reference solutions and numerical solutions for the Riemann problem RP1 at $t = 1.0$ .	127
4.22	Reference solutions and numerical solutions for the Riemann problem RP2 at $t = 1.0$ .	127
4.23	Reference solutions and numerical solutions for the Riemann problem RP3 at $t = 1.0$ .	128
4.24	Reference solutions and numerical solutions for the Riemann problem RP4 at $t = 1.0$ .	128
4.25	Reference solutions and numerical solutions for the Riemann problem RP5 at $t = 1.0$ .	129
4.26	Reference solutions and numerical solutions for the Riemann problem RP6 at $t = 1.0$ .	129
4.27	Reference solutions and numerical solutions for the Riemann problem RP7 at $t = 1.0$ .	130
4.28	Reference solutions and numerical solutions for the Riemann problem RP8 at $t = 0.75$ .	130
4.29	Reference solutions and numerical solutions for the Riemann problem RP9 at $t = 0.7$ .	131
4.30	Reference solution and numerical solution for the trans-critical flow over a bump at $t = 1.0$ .	133
4.31	Oscillating lake - Reference solutions and numerical solutions for $\eta$ and $Q$ at $t = \tau/4$ .	134
4.32	Oscillating lake - Reference solutions and numerical solutions for $\eta$ and $Q$ at $t = \tau/2$ .	134
4.33	Oscillating lake - Reference solutions and numerical solutions for $\eta$ and $Q$ at $t = 3\tau/4$ .	135
4.34	Oscillating lake - Reference solutions and numerical solutions for $\eta$ and $Q$ at $t = \tau$ .	135
4.35	Numerical validation of the well-balanced property for the 2D case. Numerical solution of the free surface elevation $\eta$ at times $t = 0.12, t = 0.24, t = 0.36$ and $t = 0.48$ from top left to bottom right.	137
4.36	Circular dam break problem with flat bottom. Contour plot of the free surface (Top left) and map of the troubled cells coloured in red (Top right); 1D cut along the $x$ axis comparing the reference solution with the numerical solution for the free surface $\eta$ (Bottom left) and for the velocity $u$ (Bottom right).	138
4.37	Circular dam break problem with discontinuous bottom. Contour plot of the free surface (Top left) and map of the troubled cells coloured in red (Top right); 1D cut along the $x$ axis comparing the reference solution with the numerical solution for the free surface $\eta$ (Bottom left) and for the velocity $u$ (Bottom right).	139
5.1	Staggered grids for the one-dimensional DG scheme. Main grid used for the free surface (top) and dual mesh for the velocity (bottom).	144
5.2	Staggered grids for the sub-cell finite volume scheme in one dimension in the case $P = 2$ ( $N_s = 5$ ). Main grid used for the free surface (top) and staggered dual mesh for the velocity (bottom).	150
5.3	Projection and reconstruction operators - Operators $\mathcal{P}$ and $\mathcal{W}$	154
5.4	Projection and reconstruction operators - Operator $\mathcal{W}_{\mathcal{L}}$	155
5.5	Projection and reconstruction operators - Operator $\mathcal{W}_{\mathcal{R}}$	156

5.6	Staggered grids for the semi-implicit staggered DG scheme with sub-cell limiter active in cell $T_i$ . Main grid used for the free surface (top) and staggered dual mesh for the velocity (bottom). If a cell $T_i$ on the main grid is flagged as troubled, then also the two overlapping staggered velocity control volumes $T_{i\pm\frac{1}{2}}$ are flagged as troubled. The data representation in troubled cells is changed from high order polynomials to piecewise constant subcell averages. . . . .	157
5.7	Control volumes for the energy equation for the staggered semi-implicit DG scheme in the limited case using $P = 2$ , i.e. $N_s = 2P + 1 = 5$ . DG method used in $T_i$ and $T_{i-\frac{1}{2}}$ ; finite volumes in $T_{i+\frac{1}{2}}$ . . . . .	158
5.8	Control volumes for the energy equation for the staggered semi-implicit DG scheme in the limited case using $P = 2$ , i.e. $N_s = 2P + 1 = 5$ . DG method used in $T_i$ and $T_{i+\frac{1}{2}}$ ; finite volumes in $T_{i-\frac{1}{2}}$ . . . . .	159
5.9	Control volumes for the momentum equation for the staggered semi-implicit DG scheme in the limited case using $P = 2$ ( $N_s = 5$ ). DG in $T_i$ ; FV in $T_{i+1}$ and in $T_{i+\frac{1}{2}}$ . . . . .	160
5.10	Control volumes for the momentum equation for the staggered semi-implicit DG scheme in the limited case using $P = 2$ ( $N_s = 5$ ). DG in $T_{i+1}$ ; FV in $T_i$ and in $T_{i+\frac{1}{2}}$ . . . . .	161
5.11	Reference solutions and numerical solutions for the problem at $t = 1$ . . . . .	164
5.12	RP1 - Reference solutions and numerical solutions for $\rho$ , $U$ and $p$ at $t_{end}$ . . . . .	166
5.13	RP2 - Reference solutions and numerical solutions for $\rho$ , $U$ and $p$ at $t_{end}$ . . . . .	167
5.14	RP3 - Reference solutions and numerical solutions for $\rho$ , $U$ and $p$ at $t_{end}$ . . . . .	168
5.15	RP4 - Reference solutions and numerical solutions for $\rho$ , $U$ and $p$ at $t_{end}$ . . . . .	169
5.16	RP5 - Reference solutions and numerical solutions for $\rho$ , $U$ and $p$ at $t_{end}$ . . . . .	170
5.17	RP6 - Reference solutions and numerical solutions for $\rho$ , $U$ and $p$ at $t_{end}$ . . . . .	171
5.18	RP7 - Reference solutions and numerical solutions for $\rho$ , $U$ and $p$ at $t_{end}$ . . . . .	172
5.19	Isentropic vortex - Contour plots for the density and the pressure at time $t = 10$ . . . . .	175
5.20	Advection of a smooth density bell in 2D - Density for $P = 5$ (high order DG) and for $P = 0$ (first order FV) at time $t = 1$ . . . . .	176
5.21	Advection of a smooth density bell in 2D - Reference solutions and numerical solutions along the $x$ axis for $\rho$ , $U$ and $p$ at time $t = 1$ . . . . .	177
5.22	Two-dimensional smooth acoustic wave propagation - Contour plots for the density and the pressure at time $t = 1$ . . . . .	178
5.23	Two-dimensional smooth acoustic wave propagation - Reference solutions and numerical solutions along the $x$ axis for $\rho$ , $U$ and $p$ at time $t = 1$ . . . . .	179
5.24	Circular explosion - Contour plots for the density and map of the troubled zones (red cell) at time $t = 0.2$ . . . . .	180
5.25	Circular explosion - Reference solutions and cut of the numerical solutions along the $x$ axis for $\rho$ , $U$ and $p$ at time $t = 0.2$ . . . . .	181
5.26	Two dimensional Riemann problem RP2D C4 - Contour plots for the density and map of the troubled zones (red cell) at time $t = 0.2$ . . . . .	183
5.27	Two dimensional Riemann problem RP2D C7 - Contour plots for the density and map of the troubled zones (red cell) at time $t = 0.2$ . . . . .	183
5.28	Two dimensional Riemann problem RP2D C8 - Contour plots for the density and map of the troubled zones (red cell) at time $t = 0.2$ . . . . .	184



5.29 Two dimensional Riemann problem RP2D C16 - Contour plots for the density and map of the troubled zones (red cell) at time $t = 0.2$ . . . . .	184
--	-----